

# Multiplicatively Interacting Factors Selectively Influencing Parameters in Multiple Response Class Processing and Rate Trees

Richard Schweickert<sup>a,\*</sup>, Zhuangzhuang Xi<sup>a</sup>

<sup>a</sup>*Department of Psychological Sciences, Purdue University, United States*

---

## Abstract

Evidence in many experiments indicates that the processes involved in producing responses are arranged in a tree structure. Evidence often indicates further that an experimental factor, such as item similarity, changes a single parameter, leaving others invariant. In typical studies, a few tree structures are hypothesized a priori, and tested by goodness of fit. With the method of Tree Inference, a tree is constructed by examining the data to see if patterns occur that are predicted when two factors selectively influence different processes (Schweickert & Chen, 2008). The patterns can reveal, for example, whether selectively influenced processes are executed in order, and what the order is. If the patterns do not occur, one can conclude that no tree is possible in which the factors selectively influence processes. In earlier work, three restrictions were imposed on the trees considered: There were two classes of responses; parameters were probabilities, bounded above by 1; and factors were assumed to change parameters associated with children of a single vertex. More general results are derived here, removing these restrictions. Results on representation, uniqueness of parameters, uniqueness of tree structure, and mixtures of trees are presented.

**Keywords:** multinomial processing trees, multiplicative factors method, selective influence, tree inference

---

---

\*Corresponding address: Department of Psychological Sciences, 365A Peirce Hall, Purdue University, 703 Third Street, Purdue University, West Lafayette, Indiana 47907, United States.  
Email address: [swike@psych.purdue.edu](mailto:swike@psych.purdue.edu) (Richard Schweickert)

## 1. Introduction

Processing tree models are used for measurement of cognitive processes in tasks in various domains of psychological research; for reviews see Batchelder and Riefer (1999) and Erdfelder, Auer, Hilbig, Assfalg, Moshagen and Nadarevic (2009). A simple processing tree is in Figure 1.

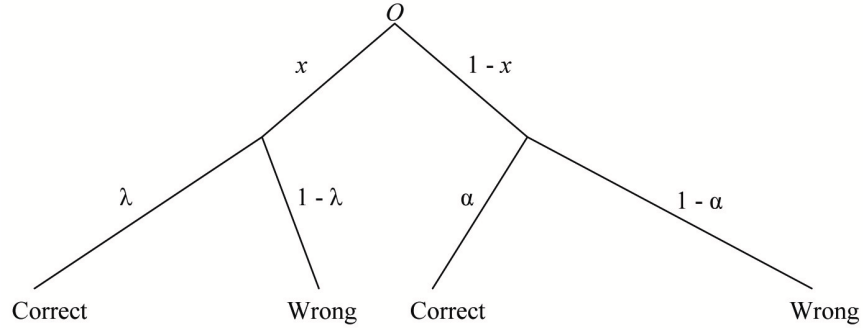


Figure 1: A processing tree with two classes, “Correct” and “Wrong”

In a *processing tree*, each vertex represents a process or a response. Processing starts at a vertex called the root ( $O$  in Figure 1). An arc descending from a vertex is called a child of the vertex; it corresponds to an outcome of the process represented by the vertex. On a given trial, a single one of these outcomes occurs; the probability of the outcome is a parameter associated with the corresponding arc. The sum of the probabilities of all the children of a vertex is 1. When an outcome occurs, we say the arc corresponding to it is *traversed*, and the ending vertex of this arc is *reached*. The process for this next vertex is then carried out. These steps continue until a vertex with no arcs descending from it is reached. Such a vertex is called a *terminal vertex*. When a terminal vertex is reached, the response corresponding to it is made.

On each trial, a response is made via a path starting from the root and reaching a terminal vertex, and the probability of the path is the product of the probabilities associated with all arcs along the path. Responses fall into various classes, and more than one path can terminate in the same class. The probability of an observed response

in a particular class is the sum of the probabilities of all the paths from the root to the terminal vertices associated with the class.

Take the processing tree in Figure 1 for example. Suppose the experimenter classifies the subject’s responses into the two classes, “Correct” and “Wrong”. The probability of a correct response equals  $x\lambda + (1 - x)\alpha$ . The parameters  $x$ ,  $\lambda$  and  $\alpha$  are probabilities and thus between 0 and 1, inclusive.

Each arc in a processing tree is associated with a probability, and the value of the probability may be determined by the level of a particular experimental factor. In other words, the probability parameters may vary with different experimental manipulations, although some may be constants. We follow Hu and Batchelder (1994) and reserve the term Multinomial Processing Tree for a processing tree for which there is a set of probabilities  $\Theta$  and each arc parameter is either an element  $\theta$  of  $\Theta$  or  $1 - \theta$ .

Considerable information is revealed about a cognitive architecture when an experimental factor can be found that changes one process or parameter, leaving others invariant, i.e., a factor that selectively influences the process or parameter. See Thomas (2006) for recent work on selective influence of parameters, and Dzhafarov (2003) and Kujala and Dzhafarov (2008) for fundamental theory. Most factors manipulated in experiments change values of several parameters, but a large body of evidence shows that certain factors selectively influence probability parameters in a processing tree, see, e.g., Riefer and Batchelder (1995) and Jacoby, Begg and Toth (1997).

A good example of selective influence in a Multinomial Processing Tree (MPT) is provided by Stahl and Degner (2007). They introduced an MPT model (the ABC model) for the Extrinsic Affective Simon Task (De Houwer, 2003). The model fit data from four experiments well. Further, the experiments validated the parameters in the model; that is, experimental manipulations expected to change a certain single parameter indeed did so.

The Extrinsic Affective Simon Task is a variant of the affective Simon task. In a typical affective Simon task, per se, stimuli are words of positive, negative or neutral valence (e.g., HAPPY). Subjects respond to each stimulus by pronouncing “positive” or “negative.” But the required response is not determined by the valence itself (i.e., whether the meaning of the word is positive or negative). Instead, the required response

is determined by a task-relevant stimulus feature such as the color the word is presented in (De Houwer & Eelen, 1998). Performance is better on trials in which the required response is congruent with the valence. In the Extrinsic Affective Simon Task (EAST), subjects are instructed to respond to the target stimulus (e.g., a positive or negative word presented in blue or green) by pressing a left or a right key based on the color. In control trials the task-relevant feature is not available (e.g., words are presented in white). On these control trials, the subject is to respond by pressing the left key if the word is positive in value and the right key if the word is negative. (Keys are counterbalanced.) As with the ordinary affective Simon Task, there is an advantage for congruent trials.

The proposed ABC model (Stahl & Degner, 2007) assumes that EAST performance is the combination of three successive cognitive processes: Automatic activation of valence ( $A$ ), controlled processing of the task-relevant feature ( $C$ ) and guessing ( $B$ ). Suppose a word with positive valence is presented in blue. With probability  $A$  the valence of the word is automatically activated, and a key is pressed based on the valence. If the valence and required key happen to be congruent the response is correct, otherwise it is incorrect. With probability  $1 - A$  the valence is not activated. Then with probability  $C$  the color (blue in this example) determines the response. If the color determines the response, it is correct. Finally, with probability  $(1 - A)(1 - C)$  the valence is not activated and the color does not determine the response. In that case, the subject guesses, and presses the left key with probability  $B$ . If the stimulus is a control stimulus presented in white, processing is the same, except that the controlled process (corresponding to probability  $C$ ) is not used.

The model predicts that a difference in automatic activation of stimulus is reflected only in a difference in the  $A$  parameter, a difference in controlled task processing is reflected only in a difference in the  $C$  parameter, and a difference in response bias is reflected only in a difference in the  $B$  parameter.

There are several examples in experiments of Stahl and Degner (2007) of factors selectively influencing processes. In Experiment 1, the similarity between the green and blue colors of the words was either high or low. Based on statistical tests, this manipulation changed the value of parameter  $C$ , the parameter corresponding to con-

trolled processing, but left the other two parameters invariant. In Experiment 4, words were presented with flankers rather than in colors. The flanker # on both sides of a word indicated one key, say, left, while the flanker @ on both sides indicated the other key. The response to a word with no flankers was indicated by the positive or negative valence of its meaning. Words of neutral valence were also presented, always with flankers. The proportion of neutral words whose flankers indicated the left key was manipulated, either 41.7% or 58.3%. Factors selectively influenced processes. Parameter  $A$  had different values for positive, negative and neutral words, but parameters  $B$  and  $C$  did not change with valence. Further, parameter  $B$  differed depending on the left key proportion, but parameters  $A$  and  $C$  did not. See their paper for details.

As another example, an MPT model for recall of an item during immediate serial recall was proposed by Schweickert (1993). The model is in Figure 2. There are two possible routes to a correct response. The trace of an item is either intact, with probability  $I$ , or degraded. If the trace is intact, the item is recalled correctly directly from memory; if the trace is degraded, a second process, redintegration, occurs, which might successfully reconstruct the item with the help of information from long-term storage. If the trace is degraded, the probability the item is redintegrated is  $R$ . The probability of a correct recall is thus  $I + (1 - I)R$ .

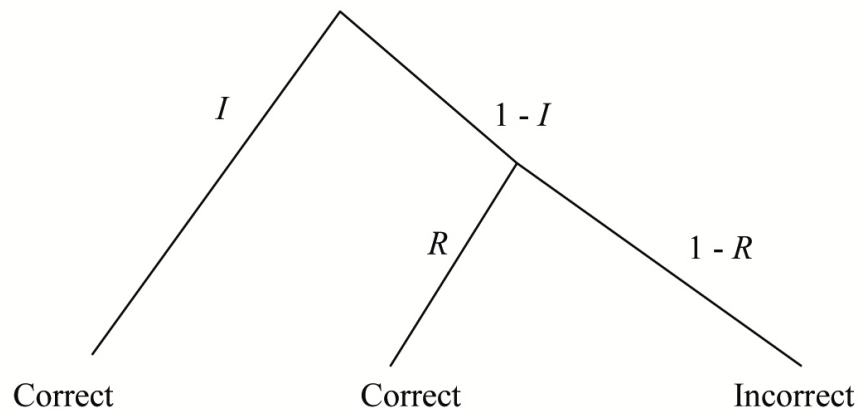


Figure 2: Schweickert's (1993) multinomial processing tree model of redintegration in immediate verbal recall

The two processes have been found to be selectively influenced by different factors. Li, Schweickert and Gandour (2000) found that a model in which a change in serial position changed the value of  $I$ , while a change in phonological similarity changed  $R$  gave a good account of the data from an immediate serial recall experiment. Another example is an experiment by Buchner and Erdfelder (2005) in which subjects memorized target words while ignoring distractors, which were frequent or rare words presented as speech. In immediate serial recall, more errors were made when distractors were high frequency than low. But distractor frequency selectively influenced the probability of intact target word representations, that is, changed the  $I$  parameter only.

### 1.1. Previous work on Tree Inference

In processing tree studies, investigators usually begin with a hypothetical model, estimate parameters from data and evaluate the model. In contrast, Schweickert and Chen (2008) provided a data-driven approach showing how to systematically construct a processing tree using data from an experiment with two factors. The approach is called *Tree Inference*. Suppose two experimental factors selectively influence two processes in a processing tree, where each process is represented by one vertex in the tree. There are only two possible arrangements for the two processes, ordered or unordered. The two processes are *ordered* if there is a path from the root to a terminal vertex that goes through both vertices; otherwise the processes are unordered. Schweickert and Chen (2008) showed that only two standard trees need be considered, one for each arrangement. If neither standard tree is possible, then no tree is possible in which each factor selectively influences a different process. For each arrangement, theorems on representation, uniqueness of parameter values and uniqueness of structure were stated, with necessary and sufficient conditions for the standard trees to be applicable.

The purpose here is to continue the development of Tree Inference by generalizing the trees considered by Schweickert and Chen (2008). The processing trees they discussed are restricted in three ways. First, only two response classes are allowed. Second, each experimental factor changes parameters associated with arcs descending from a single vertex. Third, parameters are assumed to be probabilities, that is, to take on values between 0 and 1.

Clearly in many cases, it is not sufficient to have only two classes. For example, MPT models are used for source monitoring experiments (e.g., Meiser & Bröder, 2002; Meiser, 2005; Dodson, Holland, & Shimamura, 1998). In a typical source monitoring experiment, subjects learn lists presented by different sources (by visual and auditory channels for example). Subjects are then presented with test lists and judge whether test items appeared in the learning phase or not, and indicate the source by which the items were presented if they were presented. A part of one of the multinomial tree models of source memory by Batchelder and Reifer (1990) is in Figure 3. Two sources are investigated, and thus there are three classes of responses (“Source A”, “Source B” and “New”).

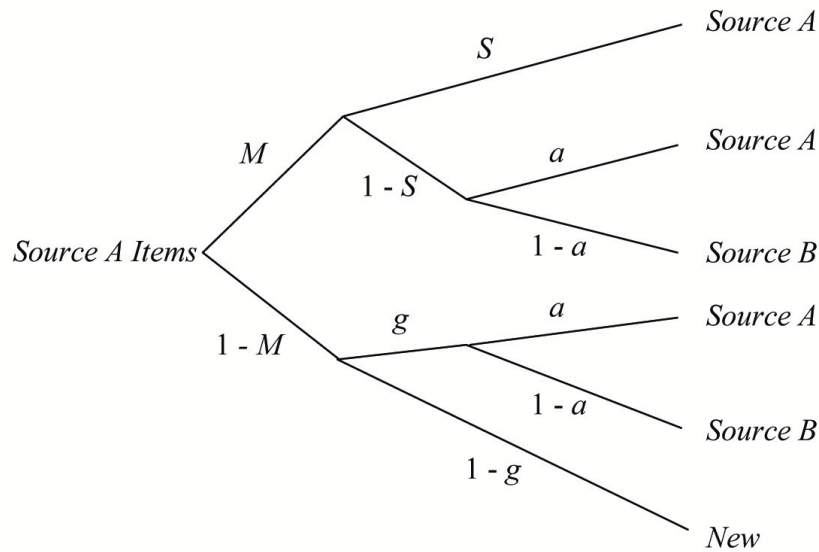


Figure 3: Part of Batchelder and Reifer’s (1990) multinomial tree model of source memory. Items are presented from two possible sources, A and B.  $M$  is the probability of correctly identifying an item as old;  $S$  is the probability of correctly identifying an item as a Source A item;  $g$  is the probability of guessing that an item is old;  $a$  is the probability of guessing that an item is from Source A

As another example, a critique of Schweickert’s (1993) model with reintegration is that it does not allow different amounts of degradation (Roodenrys & Miller, 2008), so Thorn, Gathercole and Frankish (2005) argue that the model should be modified to

allow three response classes, “completely correct,” “partially correct” and “completely wrong”. The model with three response classes gives a better account of their data than the model with two. For recent discussions of factors important for immediate memory, see Thorn and Page (2009).

Regarding the second restriction, it is not uncommon for the same parameter to occur at more than one vertex in a processing tree. In the part of the Batchelder and Riefer (1990) MPT model for source monitoring in Figure 3, parameter  $a$  is the probability of guessing that an item was presented by Source A. Two vertices are associated with the  $a$  parameter. If a factor changes the value of  $a$  it would change parameters at more than one vertex.

A third restriction in the theorems of Schweickert and Chen (2008) is that parameters are probabilities, taking values between 0 and 1. But trees whose parameters are not probabilities are often used in psychology, with evidence for experimental factors selectively influencing processes. In particular, Roberts (1987) used trees to model response rates of behaviors, such as lever pressing, in animal experiments. In one of his models, a pulse generator produces pulses at a rate  $g$  and a filter sends a proportion  $p$  of the pulses to the response system. Parameter  $p$  is between 0 and 1, but parameter  $g$  is not. The model was successfully applied to an experiment of Clark (1958), in which rats pressed a lever for food. There were three reward schedules, Variable Intervals of 1 min, 2 min, and 3 min. There were seven delays of testing after eating, 1, 3, 5, 7, 10, 20 and 23 hours. The two factors had multiplicative effects on the rate of lever pressing at test. In the model, the pulse rate  $g(i)$  depends on the level  $i$  of delay and the proportion  $p(j)$  of pulses sent to the response system depends on the level  $j$  of reward schedule. The rate of lever pressing for factor combination  $(i, j)$  is

$$r(i, j) = g(i) \times p(j).$$

The factors have multiplicative effects, the basis for the Multiplicative Factors Method of Roberts (1987). He found considerable evidence for multiplicative factors in rate experiments with rats, pigeons and goldfish.



### 1.2. The standard tree for multiplicatively interacting factors

Here theorems are generalized to processing trees with more than two response classes, in which two experimental factors have multiplicatively interacting effects. A processing tree with such properties is in Figure 4 for two factors selectively influencing two different sets of parameters. When one factor is at level  $i$  and the other is at level  $j$ , the probability of the  $k$ th response  $R_k$  is

$$Prob[R_k] = w_{i,k} + x_i y_{j,k} + z_{j,k}. \quad (1)$$

**Definition.** The *standard tree for multiplicatively interacting factors* has a root vertex with three types of children. There are  $C$  children of one type, each associated with a probability  $w_{i,k}$ , and each ending at a terminal vertex for a different response class. There are  $C$  children of another type, each associated with a probability  $z_{j,k}$ , and each ending at a terminal vertex for a different response class. There is 1 child of the third type, associated with a probability  $x_i$ ; it has  $C$  children, each associated with a probability  $y_{j,k}$ , and each ending at a terminal vertex for a different vertex.

In some papers, the term “tree” refers to a graph without parameters. Here when we use the term “tree” we refer to a graph (vertices and arcs) together with the parameters associated with its arcs.

**Definition.** By saying the two factors *selectively influence two different sets of parameters* in a processing or rate tree we mean the following. Parameters are written so those in one subset have an index including an  $i$  and those in another subset, mutually exclusive from the first, have an index including a  $j$ . (Some parameters may not have an index with an  $i$  or a  $j$ .) Changing the level  $i$  of one factor changes only parameters whose index includes  $i$  and changing the level  $j$  of the other factor changes only parameters whose index includes  $j$ .

In this paper, when we say two factors selectively influence parameters in the standard tree for multiplicatively interacting factors, we mean the parameters are written as in Equation (1), so changing the level  $i$  of one factor changes only parameters  $w_{i,k}$  and  $x_i$ , and changing the level  $j$  of the other factor changes only parameters  $y_{j,k}$  and  $z_{j,k}$ . Because the sum of the probabilities associated with the children of the root is always 1, at the root of the tree the sum over  $k$  of the parameters  $z_{j,k}$  must be a constant,  $b$ ,

and the sum over  $k$  of all the parameters  $y_{j,k}$  must always be 1. The parameter  $b$  is not explicitly associated with an arc in the tree, it is implicit in the assumption of the selective influence of the factors. Only one arc with a parameter indexed by  $i$  is followed by arcs with parameters indexed by  $j$ . This leads to interactions that are products.

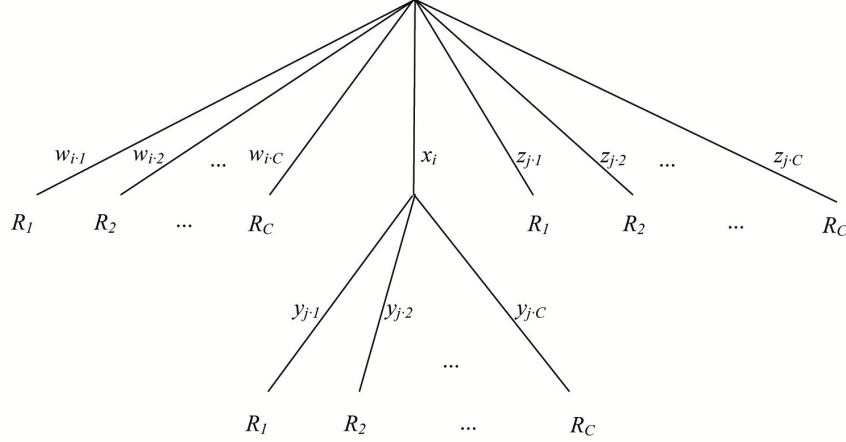


Figure 4: The standard tree for multiplicatively interacting factors, with an arc associated with  $i$  preceding arcs associated with  $j$

Two processing trees are *equivalent* for two factors if the probability of every response class is the same for each tree at every combination of levels of the factors. In the standard tree for multiplicatively interacting factors, parameters indexed by  $i$  occur at the same vertex as parameters indexed by  $j$ . There is an equivalent tree in which each vertex has children associated with parameters indexed by either  $i$  or by  $j$  but not both, see Figure 5. In this tree the parameter  $b$  is explicit, at the cost of additional arcs.

### 1.3. Examples

#### Example 1. Weapon Misidentification

When subjects are shown a picture of a face followed by a picture of a weapon or a tool (e.g., a gun or a drill), they are more likely to misidentify the tool as a weapon if a Black face was presented than a White face (Payne, 2001). For the experimental situation, Bishara and Payne (2009) compare two models, which they call the Process Dissociation Model and the Stroop Model. The two models are similar; they make

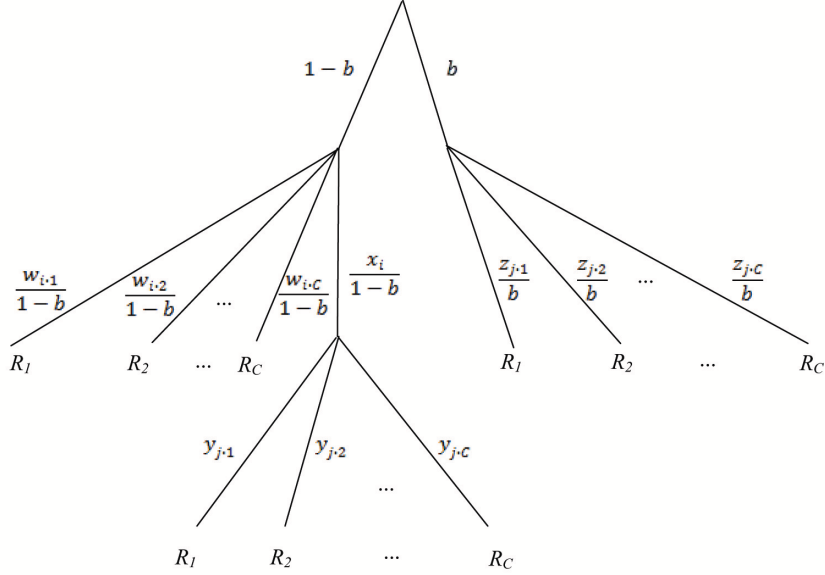


Figure 5: An equivalent standard tree for multiplicatively interacting factors, with an arc associated with  $i$  preceding arcs associated with  $j$

identical predictions for the White Tool and Black Gun conditions, but different predictions for the White Gun and Black Tool conditions. Bishara and Payne show that both models can be nested in a single model, the Quad Model of Conrey, Sherman, Gawronski, Hugenberg and Groom (2005). It is in Figure 6; the left hand arc from the root leads to the Stroop with Guessing Model, the right hand arc to the Process Dissociation with Guessing Model. After fitting the Quad Model to their data and those of others, Bishara and Payne (2001) conclude there is more support for the Process Dissociation Guessing Model than for the Stroop with Guessing Model.

Bishara and Payne (2001) discuss different ways of writing the models and interpreting the parameters; details are beyond the scope of our paper. Briefly,  $OB$  can be interpreted as the probability of following the Process Dissociation with Guessing Model,  $1 - OB$  as the probability of following the Stroop with Guessing Model. In either model, with probability  $C$  controlled processing succeeds, always leading to a correct identification. With probability  $A$  automatic influence occurs, leading to a tool

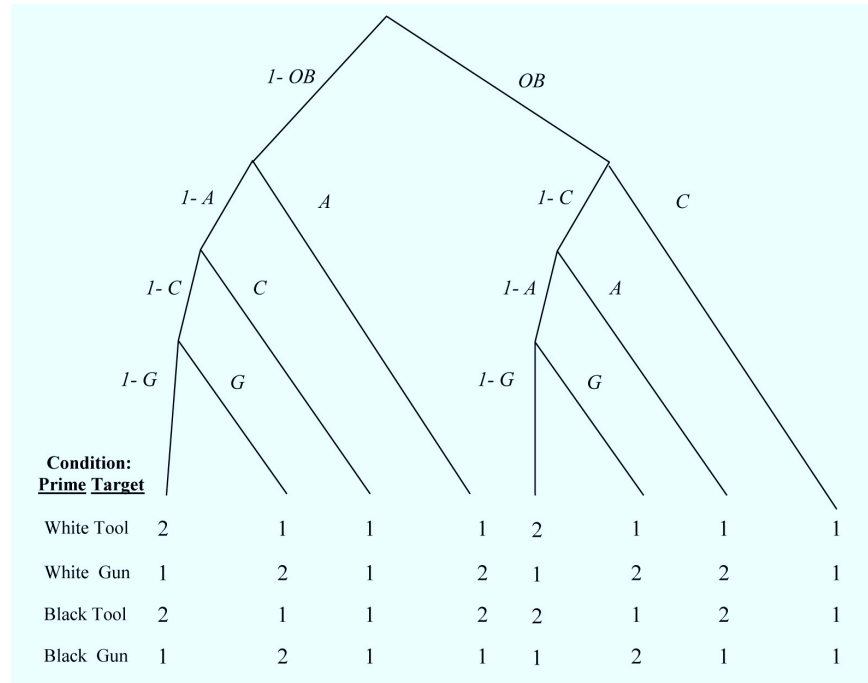


Figure 6: The Quad Model as in Bishara and Payne (2009). Response 1 is correct, 2 is wrong

response for a White face and a gun response for a Black face. If neither controlled processing nor automatic influence occurs, with probability  $G$  the subject guesses, leading to a tool response. Cornrey, et al. (2005), using the Quad Model in a different task, the Implicit Association Test (Greenwald, McGhee, & Schwartz, 1998), found factors that selectively influence parameters. In particular, a change in association from insect-unpleasant to flower-pleasant changed only parameter  $A$ , while changing the frequency of correct responses requiring the right hand changed only parameter  $G$ .

In the Quad Model (Figure 6), parameter  $A$  occurs in more than one place, so one might at first think an experimental factor changing the value of  $A$  would have untidy effects. Parameter  $G$  also occurs at more than one place in the tree. But suppose when one factor is at level  $i$ , the value of  $A$  is  $A_i$ , and when another factor is at level  $j$ , the

value of  $G$  is  $G_j$ . For the White Gun condition,

$$p_{ij \cdot 1} = OB * C + (1 - OB)C(1 - A_i) + (1 - C)(1 - A_i)(1 - G_j) \quad (2)$$

$$p_{ij \cdot 2} = (1 - OB * C)A_i + (1 - C)(1 - A_i)G_j. \quad (3)$$

The form of these equations shows the Quad Model for the White Gun condition is equivalent to the standard tree with multiplicative interaction. The same is true for the other conditions, and analysis is straightforward.

*Example 2. Source Monitoring*

A multinomial processing tree that can sometimes take the form of the standard tree in Figure 4 is the part of the source memory model of Batchelder and Reifer (1990) in Figure 3. Suppose the form of the underlying tree is unknown, but we manipulate a factor that changes parameter  $M$  and another factor that changes parameter  $a$ . When the first factor is at level  $i$ , let the value of  $M$  be  $M_i$  and when the second factor is at level  $j$ , let the value of  $a$  be  $a_j$ . It is easily seen that

$$p_{ij \cdot A} = M_i S + M_i a_j (1 - S - g) + g a_j;$$

$$p_{ij \cdot B} = M_i (1 - a_j) (1 - S - g) + g (1 - a_j);$$

$$p_{ij \cdot \text{New}} = (1 - M_i) (1 - g).$$

Here, letting  $k$ , the index for response classes, take values  $A$ ,  $B$  and  $\text{New}$  rather than 1, 2 and 3,

$$w_{i \cdot A} = M_i S, \quad w_{i \cdot B} = 0, \quad w_{i \cdot \text{New}} = (1 - M_i) (1 - g),$$

$$x_i = M_i (1 - S - g),$$

$$y_{j \cdot A} = a_j, \quad y_{j \cdot B} = (1 - a_j), \quad y_{j \cdot \text{New}} = 0, \quad (4)$$

$$z_{j \cdot A} = g a_j, \quad z_{j \cdot B} = g (1 - a_j), \quad z_{j \cdot \text{New}} = 0.$$

If probabilities  $S$  and  $g$  have values such that

$$1 - S - g \geq 0,$$

the model is in the form of the standard tree for multiplicatively interacting factors. (If the inequality above is violated, parameter  $x_i$  becomes negative and cannot be a probability.) Note that  $b$  in Figure 5 is equal to  $g$  here. That is,  $z_{j \cdot A} + z_{j \cdot B} + z_{j \cdot \text{New}} = g$ .

*Example 3. Rates*

The most complex rate model considered by Roberts (1987), his Model 3, has the form of the standard tree in Figure 5. His model has the following equation, for the rate of responding when one factor has level  $i$  and another factor has level  $j$ ,

$$r(i, j) = c(i) + a(i) \times b(j) + d(j). \quad (5)$$

Here the model (5) is called the single-response rate model for two multiplicatively interacting factors. It is assumed that a mechanism generates an animal behavior, such as lever pressing, through three distinct paths. One path, the main path, consists of two ordered processes, a pulse generator and a filter, which are selectively influenced by two different factors. The first process generates pulses, which travel to a filter, and pulses that pass through the filter produce a response. The response rate produced through this path is the product of pulse-generation rate,  $a(i)$ , and the probability that the pulse is transmitted by the filter,  $b(j)$ . A second path produces responses at rate  $c(i)$ , which is selectively influenced by one of the factors, and a third path produces responses at rate  $d(j)$ , which is selectively influenced by the other factor.

It is natural to generalize the single-response rate model for two multiplicatively interacting factors to allow two or more types of responses, when, for example, animals make different responses to different stimuli, or a conductor moves a baton with one hand and gestures with the other. It is also natural to require all parameters to be nonnegative, but not to restrict them to be less than 1. For want of a better term, we call a tree whose parameters are nonnegative a *rate tree*, although the parameters need not be rates. A standard rate tree in which multiple response classes are allowed is in Figure 7. In the model the  $k$ th class of response  $R_k$  can be made through three paths: the main path consists of a pulse generator and a filter, where the rate of the pulse generator is selectively influenced by Factor 1 and the filter probability is selectively influenced by Factor 2; the second and third paths consist of response-specific pulse generators whose rates are selectively influenced by Factor 1 and Factor 2, respectively. In order to distinguish between different pulse generators, the one followed by a filter is called the central pulse generator. In this model, the rate of the central pulse generator is the same for all response types. For convenience of comparison between rate models

and processing tree models, the same parameter notation will be used. The rate for response  $R_k$  is

$$Rate[R_k] = w_{i \cdot k} + x_i y_{j \cdot k} + z_{j \cdot k}.$$

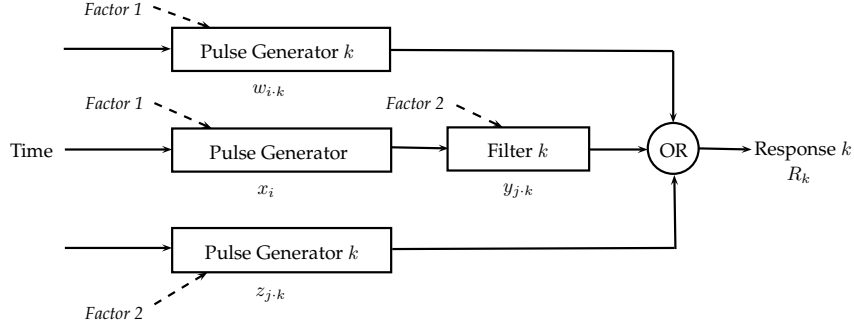


Figure 7: The multiple-responses rate model for two multiplicatively interacting factors, with the central pulse generator indexed by  $i$

Theorems are developed here on representation, uniqueness of parameter values and uniqueness of tree structure. Statistical issues are important, but beyond the scope of this paper; we assume data are free of noise.

## 2. Selective Influence of Two Multiplicatively Interacting Factors in The Generalized Processing Tree Models

### 2.1. Notation

Before theorems are stated notation is introduced.

#### 2.1.1. C-class processing and rate trees.

Suppose an experiment has two factors, Factor 1 and Factor 2, which change parameters associated with some arcs in a processing or rate tree, leaving other parameters invariant. If no parameter is changed by both factors, each factor is said to *selectively influence* the parameters it changes. Suppose Factor 1 has levels  $i = 1, 2, \dots, I$  and Factor 2 has levels  $j = 1, 2, \dots, J$ . The tree is called a *two-indexed tree*. In this paper, we only consider two-indexed trees.

When there are two response classes, say coded as “correct” and “wrong”, responses produced via a processing tree can be represented in a probability matrix, which is defined here as a matrix whose entries are numbers between 0 and 1, inclusive. (The term probability matrix sometimes denotes a matrix of nonnegative numbers with the sum of each row being 1. That is not the meaning here.) We have an  $I \times J$  probability matrix,

$$P_{I \times J} = \begin{pmatrix} p_{11} & p_{12} & \dots & p_{1J} \\ p_{21} & p_{22} & \dots & p_{2J} \\ \dots & \dots & \dots & \dots \\ p_{I1} & p_{I2} & \dots & p_{IJ} \end{pmatrix}.$$

The cell value  $p_{ij}$  is the probability of a correct response when Factor 1 is at level  $i$  and Factor 2 is at level  $j$ . A processing tree with two response classes is called a two-class processing tree.

If there are more than two classes, say  $C$  classes (where  $C$  is a positive integer), the processing tree is called a  $C$ -class processing tree. In a  $C$ -class processing tree, a response class is denoted  $R_k$ , where  $k = 1, 2, \dots, C$ . For each class, a probability matrix can be defined in the same way as for a two-class processing tree. However, to present all the data, a probability matrix stack is required. A probability matrix stack is a set of probability matrices, each layer being the probability matrix corresponding to a particular class of response. To be specific, for a particular response class  $R_k$  in a  $C$ -class processing tree, there is a probability matrix  $P_{(k)} = (p_{ij \cdot k})_{I \times J}$ , where  $p_{ij \cdot k}$  is the probability of the  $k$ th response class  $R_k$  when Factor 1 is at level  $i$  and Factor 2 is at level  $j$ .

$$P_{(k)} = (p_{ij \cdot k})_{I \times J} = \begin{pmatrix} p_{11 \cdot k} & p_{12 \cdot k} & \dots & p_{1J \cdot k} \\ p_{21 \cdot k} & p_{22 \cdot k} & \dots & p_{2J \cdot k} \\ \dots & \dots & \dots & \dots \\ p_{I1 \cdot k} & p_{I2 \cdot k} & \dots & p_{IJ \cdot k} \end{pmatrix}.$$

In a  $C$ -class processing tree, there are  $C$  such probability matrices  $P_{(k)}$ ,  $k = 1, 2, \dots, C$ . A probability matrix stack is denoted  $\{P_{(k)}\}_C = \{P_{(1)}, P_{(2)}, \dots, P_{(C)}\}$ . Note that because exactly one of the  $C$  response classes must occur, for any pair  $(i, j)$ ,  $1 \leq i \leq I$  and  $1 \leq j \leq J$ , it must be that  $\sum_{k=1}^C p_{ij \cdot k} = 1$ .



Similarly, suppose two experimental factors selectively influence processes in a two-indexed rate tree. To represent response rates produced by the rate tree, a rate matrix stack is needed. To keep expressions simple, the same notation is used for a rate matrix stack and a probability matrix stack,  $\{P_{(k)}\}_C = \{P_{(1)}, P_{(2)}, \dots, P_{(C)}\}$ . Each layer is an  $I \times J$  matrix whose element  $p_{ij \cdot k}$  is the rate for response  $R_k$  when Factor 1 is at level  $i$  and Factor 2 is at level  $j$ . When we talk about rate models,  $p_{ij \cdot k}$  is nonnegative, but does not need to be less than or equal to 1, and it is not required that  $\sum_{k=1}^C p_{ij \cdot k} = 1$ .

### 2.1.2. *Multiplicatively interacting factors.*

The standard  $C$ -class processing tree for two multiplicatively interacting factors, with an arc associated with  $i$  preceding arcs associated with  $j$ , is in Figure 4. When Factor 1 is at level  $i$  and Factor 2 is at level  $j$ , the probability of the  $k$ th response class  $R_k$  is

$$p_{ij \cdot k} = w_{i \cdot k} + x_i y_{j \cdot k} + z_{j \cdot k}. \quad (6)$$

The two factors are multiplicatively interacting because of the term  $x_i y_{j \cdot k}$ .

The same equation applies to a standard  $C$ -class rate model for two multiplicative interacting factors, illustrated in Figure 7, with the central pulse generator indexed by  $i$ . When Factor 1 is at level  $i$  and Factor 2 is at level  $j$ , the equation gives the rate for response  $R_k$ , where  $k = 1, 2, \dots, C$ . In the rate model,  $p_{ij \cdot k}$  represents the rate of a response rather than probability, and we assume parameters  $w_{i \cdot k}$ ,  $x_i$ ,  $y_{j \cdot k}$  and  $z_{j \cdot k}$  are nonnegative, but not necessarily bounded above by 1.

For the standard processing or rate tree recall that for all  $j$ ,  $\sum_{k=1}^C z_{j \cdot k} = b$ , where  $b$  is a constant. This is needed because changing the level  $j$  of Factor 2 does not change  $w_{i \cdot k}$  or  $x_i$ .

Also, in the standard processing tree, recall that the sum of the probabilities associated with the children of a vertex is 1, so  $\sum_{k=1}^C y_{j \cdot k} = 1$ . This is not necessarily the case in the standard rate model, even if each  $y_{j \cdot k}$  happens to be bounded above by 1.

Interactions are important, of course, for describing the properties of a tree for multiplicatively interacting factors. The effect of changing Factor 1 from level  $i^*$  to level  $i$  may depend on the level  $j$  of Factor 2. When Factor 2 is at level  $j$  this effect is

$p_{ij \cdot k} - p_{i^*j \cdot k}$ . When Factor 2 is at level  $j^*$  this effect is  $p_{ij^* \cdot k} - p_{i^*j^* \cdot k}$ . If these effects are the same, the interaction is 0 for these levels. In general, for any two pairs of factor levels (i.e., two experimental conditions),  $(i, j)$  and  $(i^*, j^*)$ , an interaction contrast for class  $k$  is defined as

$$p_{ij \cdot k} - p_{i^*j \cdot k} - p_{ij^* \cdot k} + p_{i^*j^* \cdot k}.$$

When the parameters from the standard tree for multiplicatively interacting factors are inserted into the interaction contrast for class  $k$ , the interaction contrast becomes

$$(x_i - x_{i^*})(y_{j \cdot k} - y_{j^* \cdot k}).$$

One multiplicand depends only on  $i$ , the other only on  $j$  and  $k$ . Note that if  $x_i \geq x_{i^*}$  and  $y_{j \cdot k} \geq y_{j^* \cdot k}$ , then both multiplicands are nonnegative. The interaction contrast is the contribution of the change in level  $i$  and the change in level  $j$  that is different from the sum of their separate effects. Interactions in processing and rate tree models are defined in the same way.

Two factors are *effective interactively for class  $k$*  if there are at least two pairs of factor levels  $(i, j)$  and  $(i^*, j^*)$  for which the interaction contrast for class  $k$  is not 0. Two factors are *effective interactively* if there is at least one class  $k$  for which the factors are effective interactively.

For an interaction in the standard tree for multiplicatively interacting factors, only parameters  $x_i$  and  $y_{j \cdot k}$  are involved. These are multiplied, so if the factors are effective interactively, we say they are *multiplicatively interacting*. It is worth noting that with the standard tree for multiplicatively interacting factors, if the parameters are probabilities, then if two factors are effective interactively for one class  $k$  they must be also for some other class  $k'$ . For, suppose for some  $k$  there are two levels  $j$  and  $j'$  with  $y_{j \cdot k} \neq y_{j' \cdot k}$ . The sum over all classes of  $y_{j \cdot k}$  is 1; so is the sum over all classes of  $y_{j' \cdot k}$ . This is not possible unless for some other class  $k'$ ,  $y_{j \cdot k'} \neq y_{j' \cdot k'}$ .

## 2.2. Generalized Trees for Two Multiplicatively Interacting Factors

### 2.2.1. Representation.

Because we discuss special cases of a general model, we could write the following theorems in terms of the general model, and then follow each theorem with remarks

on what needs to be added to develop special cases. With this approach, however, it is cumbersome to identify the places in a proof where the additional statements need to be inserted. Instead, we write the theorems in terms of one special case, processing trees with multiplicatively interacting factors, and follow each theorem with remarks indicating what statements can be omitted to generalize.

**Theorem 1.** *Probability matrix stack  $\{P_{(k)}\}_C$  is produced by two factors effective interactively and selectively influencing parameters in the standard tree for multiplicatively interacting factors, written as in Equation (1) with an arc associated with  $i$  preceding arcs associated with  $j$ , if and only if the following three conditions are met:*

(1) *There exists a level  $i^*$  of Factor 1 and for every  $k'$  there exists a level  $j^*(k')$  of Factor 2, such that for every  $i, j$  and  $k$ , there exist  $r_i, 0 \leq r_i \leq 1$  and  $s_{j \cdot k}, 0 \leq s_{j \cdot k} \leq 1$ , such that*

$$p_{ij \cdot k} \geq p_{ij \cdot k} - p_{ij^*(k) \cdot k} - p_{i^*j \cdot k} + p_{i^*j^*(k) \cdot k} = r_i s_{j \cdot k}.$$

(2) *For two levels of Factor 1,  $i$  and  $i'$ ,  $r_i \neq r_{i'}$ . Further, for some  $k$ , for two levels  $j$  and  $j'$  of Factor 2,  $s_{j \cdot k} \neq s_{j' \cdot k}$ .*

(3) *For every  $j$ ,  $\sum_{k=1}^C s_{j \cdot k} = a$ , where  $a$  is a constant and  $0 < a \leq 1$ .*

*Proof.* (a) Suppose for every  $i, j$  and  $k$ ,

$$p_{ij \cdot k} = w_{i \cdot k} + x_i y_{j \cdot k} + z_{j \cdot k},$$

where  $0 \leq w_{i \cdot k}, x_i, y_{j \cdot k}, z_{j \cdot k} \leq 1$ , and for every  $j$ ,  $\sum_k y_{j \cdot k} = 1$  and  $\sum_k z_{j \cdot k} = b$ .

We first find the quantities specified in Condition (1). Let  $i^*$  be selected so  $x_i$  takes on its smallest value at  $i^*$ , and for every  $k$ , let  $j^*(k)$  be selected so  $y_{j \cdot k}$  takes on its smallest value at  $j^*(k)$ .

$$\text{Then } p_{ij \cdot k} - p_{ij^*(k) \cdot k} - p_{i^*j \cdot k} + p_{i^*j^*(k) \cdot k} = (x_i - x_{i^*})(y_{j \cdot k} - y_{j^*(k) \cdot k}).$$

$$\text{For every } i, j \text{ and } k, \text{ let } r_i = x_i - x_{i^*} \text{ and } s_{j \cdot k} = y_{j \cdot k} - y_{j^*(k) \cdot k}.$$

Because  $0 \leq x_{i^*} \leq x_i \leq 1$  and  $0 \leq y_{j^*(k) \cdot k} \leq y_{j \cdot k} \leq 1$ , for all  $i, j$  and  $k$ ,  $0 \leq r_i \leq 1$  and  $0 \leq s_{j \cdot k} \leq 1$ .

$$\text{Then } p_{ij \cdot k} - p_{ij^*(k) \cdot k} - p_{i^*j \cdot k} + p_{i^*j^*(k) \cdot k} = r_i s_{j \cdot k}.$$

For Condition (3),  $\sum_{k=1}^C s_{j \cdot k} = \sum_{k=1}^C y_{j \cdot k} - \sum_{k=1}^C y_{j^*(k) \cdot k} = 1 - \sum_{k=1}^C y_{j^*(k) \cdot k} = a$ , where  $a$  is a constant. Note that  $\sum_{k=1}^C y_{j^*(k) \cdot k}$  is a constant.

Clearly,  $a \leq 1$ .

Also,  $a$  is minimized at 0 if and only if  $y_{j \cdot k} = y_{j^*(k) \cdot k}$  for all  $j$  and  $k$ , that is,  $y_{j \cdot k}$  is a constant for every  $k$ . The factors are then not effective interactively, contrary to our assumption.

Therefore,  $0 < a \leq 1$ .

To complete Condition (1), because

$$\begin{aligned} & p_{ij \cdot k} - p_{ij^*(k) \cdot k} - p_{i^* j \cdot k} + p_{i^* j^*(k) \cdot k} \\ &= (x_i - x_{i^*}) \cdot (y_{j \cdot k} - y_{j^*(k) \cdot k}) \\ &\leq x_i y_{j \cdot k}, \end{aligned}$$

and  $p_{ij \cdot k} = w_{i \cdot k} + x_i y_{j \cdot k} + z_{j \cdot k} \geq x_i y_{j \cdot k}$ ,

it follows that

$$p_{ij \cdot k} \geq p_{ij \cdot k} - p_{ij^*(k) \cdot k} - p_{i^* j \cdot k} + p_{i^* j^*(k) \cdot k}.$$

Condition (2) follows immediately because the factors are effective interactively.

(b) Suppose Conditions (1), (2) and (3) hold. The proof consists of finding parameters for which Equation (1) holds.

For every  $k$ , let

$$\begin{aligned} w_{i \cdot k} &= p_{ij^*(k) \cdot k} - \min_i \{p_{ij^*(k) \cdot k}\}, \\ z_{j \cdot k} &= p_{i^* j \cdot k} + \min_i \{p_{ij^*(k) \cdot k}\} - p_{i^* j^*(k) \cdot k}, \\ x_i &= ar_i, \\ y_{j \cdot k} &= s_{j \cdot k}/a. \end{aligned}$$

From Condition (2), there are at least two different values of  $r_i$  and for some  $k$  there are at least two different values of  $y_{j \cdot k}$ .

With these parameters,

$$\begin{aligned} w_{i \cdot k} + x_i y_{j \cdot k} + z_{j \cdot k} &= p_{ij^*(k) \cdot k} - \min_i \{p_{ij^*(k) \cdot k}\} + p_{i^* j \cdot k} + \min_i \{p_{ij^*(k) \cdot k}\} \\ &\quad - p_{i^* j^*(k) \cdot k} + p_{ij \cdot k} - p_{ij^*(k) \cdot k} - p_{i^* j \cdot k} + p_{i^* j^*(k) \cdot k} \\ &= p_{ij \cdot k}. \end{aligned}$$

Clearly,  $0 \leq w_{i \cdot k} \leq 1$ .

Since  $\min_i \{p_{ij^*(k) \cdot k}\} \leq p_{i^*j^*(k) \cdot k}$ , then  $z_{j \cdot k} \leq p_{i^*j \cdot k} \leq 1$ .

According to Condition (1), for every  $k$ , for any  $i$  and  $j$ ,

$$p_{ij^*(k) \cdot k} + p_{i^*j \cdot k} - p_{i^*j^*(k) \cdot k} \geq 0.$$

For convenience, let  $i'$  be such that  $p_{i'j^*(k) \cdot k} = \min_i \{p_{ij^*(k) \cdot k}\}$ .

Then clearly,  $p_{i'j^*(k) \cdot k} + p_{i^*j \cdot k} - p_{i^*j^*(k) \cdot k} \geq 0$ , i.e.,  $z_{j \cdot k} \geq 0$ .

Therefore,  $0 \leq z_{j \cdot k} \leq 1$ .

From  $0 < a \leq 1$  and  $0 \leq r_i \leq 1$ , it follows that  $0 \leq x_i \leq 1$ .

Also, for all  $j$  and  $k$ ,  $0 \leq y_{j \cdot k} = s_{j \cdot k}/a \leq 1$ .

Further,

$$\sum_{k=1}^C y_{j \cdot k} = \sum_{k=1}^C s_{j \cdot k}/a = 1.$$

$$\sum_{k=1}^C z_{j \cdot k} = \sum_{k=1}^C p_{i^*j \cdot k} + \sum_{k=1}^C [\min_i \{p_{ij^*(k) \cdot k}\} - p_{i^*j^*(k) \cdot k}].$$

For every  $j$ ,  $\sum_{k=1}^C p_{i^*j \cdot k} = 1$ , and for every  $k$ ,  $\min_i \{p_{ij^*(k) \cdot k}\} \leq p_{i^*j^*(k) \cdot k}$ .

Let  $b^* = \sum_{k=1}^C [p_{i^*j^*(k) \cdot k} - \min_i \{p_{ij^*(k) \cdot k}\}]$ . Then  $b^*$  is a non-negative constant

which does not depend on  $j$ .

Then  $\sum_{k=1}^C z_{j \cdot k} = 1 - b^*$ .

Note that

$$\begin{aligned} \sum_{k=1}^C p_{ij \cdot k} &= \sum_{k=1}^C w_{i \cdot k} + \sum_{k=1}^C z_{j \cdot k} + \sum_{k=1}^C x_i y_{j \cdot k} \\ &= \sum_{k=1}^C w_{i \cdot k} + 1 - b^* + x_i \sum_{k=1}^C y_{j \cdot k} \\ &= \sum_{k=1}^C w_{i \cdot k} + 1 - b^* + x_i. \end{aligned}$$

Given  $\sum_{k=1}^C p_{ij \cdot k} = 1$ , then  $\sum_{k=1}^C w_{i \cdot k} + x_i = b^*$ , which is a constant independent of  $j$ . Note that  $b^*$  equals  $1 - b$  in Figure 5.

Therefore,  $\{P_{(k)}\}_C$  is produced by a standard processing tree for two multiplicatively interacting factors selectively influencing parameters, with an arc associated with  $i$  preceding arcs associated with  $j$ .  $\square$

*Remark 1.* For a  $C$ -class rate model for two multiplicatively interacting factors as in Figure 7, the representation theorem only requires slight modifications. Of course

$\{P_{(k)}\}_C$  now represents a rate matrix stack and parameters need not be bounded above by 1. The main modification is that Condition (3) is not needed because in rate models we do not assume  $\sum_{k=1}^C y_{j \cdot k} = 1$ , and thus  $\sum_{k=1}^C s_{j \cdot k}$  is not necessarily a constant but is a positive number, which may vary with  $j$ , the level of Factor 2. Further, if  $x_i$  is not bounded above by 1,  $r_i$  is not bounded above by 1 in Condition (1); similarly, if  $y_{j \cdot k}$  is not bounded above by 1,  $s_{j \cdot k}$  is only required to be non-negative. The modified proof follows directly that for Theorem 1.

*Remark 2.* If the two factors are not effective interactively, Theorem 1 only requires slight modification for either processing or rate trees. All interaction contrasts are 0. Condition (1) remains as stated, with  $r_i = 0$  for every  $i$  and  $s_{j \cdot k} = 0$  for every  $j$  and  $k$ . Condition (2) is omitted, of course. Condition (3) remains as stated, except that  $a = 0$ . We set  $x_i = r_i = 0$  for every  $i$  and we set  $y_{j \cdot k} = s_{j \cdot k} = 0$  for every  $j$  and  $k$ . In the proof, we no longer conclude that  $\sum_{k=1}^C y_{j \cdot k} = 1$  (instead the sum equals 0), but this is of no consequence.

*Remark 3.* Suppose there are parameters that produce probability matrix stack  $\{P_{(k)}\}_C$  with a standard C-class processing tree  $p_{ij \cdot k} = w_{i \cdot k} + x_i y_{j \cdot k} + z_{j \cdot k}$ . To use Theorem 1 to find parameter values, one must know the level  $i^*$  at which  $x_i$  takes its smallest value and for every class  $k$ , the level  $j^*(k)$  at which  $y_{j \cdot k}$  takes its smallest value. When the level  $i$  changes,  $w_{i \cdot k}$  might decrease while  $x_i$  increases, so the largest value of  $x_i$  may not occur at the largest value of  $p_{ij \cdot k}$ .

Values of  $x_i$  and of  $y_{j \cdot k}$  can be put in an order as follows. It turns out that two orders are possible, one the reverse of the other. If the order of the  $x_i$  parameters is reversed, then the order of the  $y_{j \cdot k}$  parameters is reversed.

According to assumption (2) of Theorem 1, there is at least one nonzero interaction contrast  $p_{i'j' \cdot k} - p_{i'j \cdot k} - p_{ij' \cdot k} + p_{ij \cdot k}$ .

Arbitrarily assume an order for  $x_{i'}$  and  $x_i$ , say  $x_{i'} > x_i$ .

Choose a level  $i^{(3)}$  of Factor 1 and form the ratio

$$\frac{p_{i^{(3)}j' \cdot k} - p_{i^{(3)}j \cdot k} - p_{ij' \cdot k} + p_{ij \cdot k}}{p_{i'j' \cdot k} - p_{i'j \cdot k} - p_{ij' \cdot k} + p_{ij \cdot k}} = \frac{x_{i^{(3)}} - x_i}{x_{i'} - x_i} \frac{y_{j' \cdot k} - y_{j \cdot k}}{y_{j' \cdot k} - y_{j \cdot k}} = \frac{x_{i^{(3)}} - x_i}{x_{i'} - x_i}.$$

If the above ratio is greater than or equal to 1, then  $x_{i^{(3)}} \geq x_i$ . Otherwise,  $x_{i^{(3)}} < x_i$ . When all such ratios are evaluated, the  $x_i$  parameters are ordered. Note that a test of

the possibility that parameters indexed by  $i$  precede parameters indexed by  $j$  is that the ratio must be the same for all classes  $k$ .

Analogous ratios can be formed for every class  $k$ , to find the order of the  $y_{j \cdot k}$  parameters.

*Remark 4.* We can find a relatively simple set of parameter values to work with.

Relabel, so  $x_1 \leq x_2 \leq \dots \leq x_I$ .

There exist parameters  $w^*$ ,  $x^*$ ,  $y^*$  and  $z^*$  such that for every  $k$ , for at least one  $i$ ,  $w_{i \cdot k}^* = 0$ . Further,  $x_1^* = 0$  and  $y_{j^*(k) \cdot k}^* = 0$ .

To see this, let  $w_{i \cdot k}^*$  be chosen as in the proof of Theorem 1, that is

$$w_{i \cdot k}^* = p_{ij^*(k) \cdot k} - \min_i \{p_{ij^*(k) \cdot k}\}.$$

Then for the value of  $i$  that minimizes  $p_{ij^*(k) \cdot k}$ ,  $w_{i \cdot k}^* = 0$ .

Let  $x_i^* = x_i - x_1$ .

Then  $x_1^* = 0$ .

Let  $y_{j \cdot k}^* = y_{j \cdot k} - y_{j^*(k) \cdot k}$ .

Then for every  $k$ ,  $y_{j^*(k) \cdot k}^* = 0$ .

Finally, let  $z_{j \cdot k}^* = z_{j \cdot k} + x_1 y_{j \cdot k} + w_{1 \cdot k} - w_{1 \cdot k}^*$ .

It is straightforward to check that

$$p_{ij \cdot k} = w_{i \cdot k}^* + x_i^* y_{j \cdot k}^* + z_{j \cdot k}^* \text{ and } 0 \leq w_{i \cdot k}^*, x_i^*, y_{j \cdot k}^*, z_{j \cdot k}^* \leq 1.$$

This raises the question of what transformations of parameters are possible, answered in the next theorem.

### 2.2.2. Uniqueness of parameters.

Suppose the standard  $C$ -class tree for multiplicatively interacting factors with a set of parameter values produces a probability matrix stack. Sometimes one set of valid parameter values can be transformed into another set of valid parameter values. An example of Buchner and Erdfelder (2005), simplified a little here, is for two response classes, say correct and wrong. Suppose when one factor is at level  $i$  and the other is at level  $j$  the probability of a correct response is  $x_i y_j$ , the product of two probabilities. There are four parameters,  $x_1, x_2$  and  $y_1, y_2$ . But the parameter values are not ordinarily unique. The product of probabilities can also be written with alternate parameters,

as  $(cx_i)(y_j/c)$  for a positive constant  $c$ . It is important to note that the value of  $c$  must be restricted because the transformed parameters must be between 0 and 1. Consequently, if the probability of a correct response is 1 for some pair of levels  $(i, j)$ ,  $c = 1$  and the parameters are unique after all. One question about uniqueness is whether the same tree with a different set of parameter values will produce the same probability matrix stack. The following theorem characterizes the possible sets of parameters that produce the same probability matrix stack.

The crux of a parameter change is the interaction term. A subtle point is that the additive terms may change in a way dependent on the probabilities involved in the interaction term. The Change of Parameter Bounds given in Table 1 are needed to keep the alternative parameter values between 0 and 1. Note that through a remark following Theorem 1 the parameters  $x_i$  can always be reindexed so  $x_1 \leq x_2 \leq \dots \leq x_I$ . With this indexing, new parameters  $x_i^*$  will be either monotonically increasing or decreasing with  $i$ .

**Theorem 2.** *Suppose  $\{P_{(k)}\}_C$  is a probability matrix stack generated by two factors that are effective interactively. Suppose probabilities  $w_{i \cdot k}$ ,  $x_i$  for  $1 \leq i \leq I$  and  $y_{j \cdot k}$ ,  $z_{j \cdot k}$  for  $1 \leq j \leq J$ ,  $1 \leq k \leq C$ , produce  $\{P_{(k)}\}_C$  through a standard tree for multiplicatively interacting factors selectively influencing parameters, with an arc associated with  $i$  preceding arcs associated with  $j$ . That is, for all  $i$  and  $j$ ,*

$$p_{ij \cdot k} = w_{i \cdot k} + x_i y_{j \cdot k} + z_{j \cdot k}.$$

*Suppose the levels  $i$  are numbered so  $x_i \geq x_{i'}$  if  $i > i'$ . Then the following two statements are equivalent.*

*(1) There exist probabilities  $w_{i \cdot k}^*$ ,  $x_i^*$  for  $1 \leq i \leq I$  and  $y_{j \cdot k}^*$ ,  $z_{j \cdot k}^*$  for  $1 \leq j \leq J$ , such that for all  $1 \leq k \leq C$ ,*

$$p_{ij \cdot k} = w_{i \cdot k}^* + x_i^* y_{j \cdot k}^* + z_{j \cdot k}^*,$$

*where the levels  $i$  are numbered so  $x_i^*$  is either monotonically increasing in  $i$  or monotonically decreasing in  $i$ . Further,  $\sum_k y_{j \cdot k}^* = 1$ .*

*(2) There exist numbers  $c \neq 0$ ;  $x_1^*$ ,  $0 \leq x_1^* < 1$ ; and for  $k = 1, \dots, C$ , numbers  $b_k$ ,*



and  $w_{1.k}^*$  such that for every  $i, j$  and  $k$ ,

$$\begin{aligned} x_i^* &= c(x_i - x_1) + x_1^* \\ y_{j.k}^* &= y_{j.k}/c + b_k \\ w_{i.k}^* &= (w_{i.k} - w_{1.k}) - b_k c(x_i - x_1) + w_{1.k}^* \\ z_{j.k}^* &= z_{j.k} + y_{j.k}(x_1 - \frac{x_1^*}{c}) - b_k x_1^* + (w_{1.k} - w_{1.k}^*) \end{aligned}$$

further, for every  $k$ , the Change of Parameter Bounds are satisfied.

*Proof.* (a) To see that Statement (2) implies Statement (1), suppose there are numbers  $c, b_k, x_1^*$  and  $w_{i.k}^*$  as described in statement (2).

Then for all  $i, j$  and  $k$ , it is easy to show

$$w_{i.k}^* + x_i^* y_{j.k}^* + z_{j.k}^* = w_{i.k} + x_i y_{j.k} + z_{j.k} = p_{ij.k}.$$

We now show the parameters are between 0 and 1.

By Statement (2),  $c \neq 0$  and  $0 \leq x_1^* < 1$ . There are two cases,  $c > 0$  and  $c < 0$ .

Suppose  $c > 0$ . The case where  $c < 0$  is similar.

From Change of Parameter Bound (18),  $c \leq \frac{1-x_1^*}{x_I-x_1}$ .

Then  $x_i^* = c(x_i - x_1) + x_1^* \leq c(x_I - x_1) + x_1^* \leq 1$ .

Also,  $x_i^* = c(x_i - x_1) + x_1^* \geq x_1^* \geq 0$ .

Thus,  $0 \leq x_i^* \leq 1$ .

Since for every  $k$ ,  $b_k c \geq -\min_j \{y_{j.k}\}$  by Change of Parameter Bound (10) and  $c(1 - b_k) \geq \max_j \{y_{j.k}\}$  by Change of Parameter Bound (16),

$$-b_k c \leq y_{j.k} \leq c(1 - b_k).$$

And thus,  $0 \leq \frac{y_{j.k}}{c} + b_k \leq 1$ , i.e.,  $0 \leq y_{j.k}^* \leq 1$ .

For  $x_i = x_1$ , Change of Parameter Bounds (19) and (20) immediately make  $0 \leq w_{i.k}^* \leq 1$ .

For  $x_i \neq x_1$ , from Change of Parameter Bounds (12) and (13),

$$\max_{x_i \neq x_1} \left\{ \frac{w_{i.k} - w_{1.k} + w_{1.k}^* - 1}{x_i - x_1} \right\} \leq b_k c \leq \min_{x_i \neq x_1} \left\{ \frac{w_{i.k} - w_{1.k} + w_{1.k}^*}{x_i - x_1} \right\}.$$

So we have

$$b_k c(x_i - x_1) \leq w_{i.k} - w_{1.k} + w_{1.k}^*$$

and

$$b_k c(x_i - x_1) \geq w_{i \cdot k} - w_{1 \cdot k} + w_{1 \cdot k}^* - 1,$$

and thus

$$0 \leq w_{i \cdot k} - w_{1 \cdot k} - b_k c(x_i - x_1) + w_{1 \cdot k}^* = w_{i \cdot k}^* \leq 1.$$

Similarly, given Inequalities (14) and (15) of Table 1, it is easy to show  $0 \leq z_{j \cdot k}^* \leq 1$ .

(b) To see that Statement (1) implies Statement (2), suppose probabilities  $w_{i \cdot k}$ ,  $x_i$ ,  $y_{j \cdot k}$  and  $z_{j \cdot k}$  and  $w_{i \cdot k}^*$ ,  $x_i^*$ ,  $y_{j \cdot k}^*$  and  $z_{j \cdot k}^*$  are such that for all  $i, j$  and  $k$ ,

$$p_{ij \cdot k} = w_{i \cdot k} + x_i y_{j \cdot k} + z_{j \cdot k} = w_{i \cdot k}^* + x_i^* y_{j \cdot k}^* + z_{j \cdot k}^*.$$

Since the factors are assumed to be effective interactively for at least one class  $k$ , we can always find a level of Factor 1, denoted  $i'$ , such that  $x_{i'} \neq x_1$ .

For every  $k$ , choose  $j'(k)$  so  $y_{j \cdot k}$  has its smallest value at  $j'(k)$ .

Consider a class  $k$  for which the factors are effective interactively. We can find a level of Factor 2,  $j''$ , such that  $y_{j'' \cdot k} \neq y_{j'(k) \cdot k}$ .

Thus, for class  $k$ ,

$$p_{i'j'' \cdot k} - p_{i'j'(k) \cdot k} - p_{1j'' \cdot k} + p_{1j'(k) \cdot k} = (x_{i'} - x_1)(y_{j'' \cdot k} - y_{j'(k) \cdot k}) \neq 0.$$

And for class  $k$ , for an arbitrary level  $i$  of Factor 1 the interaction for  $(i, j'')$  and  $(i', j'(k))$  is

$$p_{ij'' \cdot k} - p_{ij'(k) \cdot k} - p_{1j'' \cdot k} + p_{1j'(k) \cdot k} = (x_i - x_1)(y_{j'' \cdot k} - y_{j'(k) \cdot k}).$$

Therefore,

$$\frac{p_{ij'' \cdot k} - p_{ij'(k) \cdot k} - p_{1j'' \cdot k} + p_{1j'(k) \cdot k}}{p_{i'j'' \cdot k} - p_{i'j'(k) \cdot k} - p_{1j'' \cdot k} + p_{1j'(k) \cdot k}} = \frac{x_i - x_1}{x_{i'} - x_1}.$$

Also,

$$\frac{p_{ij'' \cdot k} - p_{ij'(k) \cdot k} - p_{1j'' \cdot k} + p_{1j'(k) \cdot k}}{p_{i'j'' \cdot k} - p_{i'j'(k) \cdot k} - p_{1j'' \cdot k} + p_{1j'(k) \cdot k}} = \frac{x_i^* - x_1^*}{x_{i'}^* - x_1^*}.$$

Thus,

$$\frac{x_i - x_1}{x_{i'} - x_1} = \frac{x_i^* - x_1^*}{x_{i'}^* - x_1^*}.$$

Therefore,

$$x_i^* = c(x_i - x_1) + x_1^*,$$

where  $c = \frac{x_{i'}^* - x_1^*}{x_{i'} - x_1}$ . If  $c = 0$ ,  $x_i^* = x_1^*$  for every  $i$  and the factors are not effective interactively.

$$\text{Because } (x_i - x_1)(y_{j \cdot k} - y_{j'(k) \cdot k}) = (x_i^* - x_1^*)(y_{j \cdot k}^* - y_{j'(k) \cdot k}^*),$$

$$y_{j \cdot k}^* = \frac{y_{j \cdot k}}{c} + b_k,$$

where  $b_k = y_{j \cdot k}^* - y_{j'(k) \cdot k}^*/c$ .

It follows that

$$y_{j \cdot k}^*(x_i^* - x_1^*) = y_{j \cdot k}x_i + b_k c(x_i - x_1) - y_{j \cdot k}x_1,$$

and

$$x_i^* y_{j \cdot k}^* = (x_i - x_1)y_{j \cdot k} + b_k c(x_i - x_1) + \frac{x_1 y_{j \cdot k}}{c} + b_k x_1^*.$$

Also, because for all  $i, j$  and  $k$ ,

$$p_{ij \cdot k} - p_{1j \cdot k} = w_{i \cdot k} - w_{1 \cdot k} + y_{j \cdot k}(x_i - x_1) = w_{i \cdot k}^* - w_{1 \cdot k}^* + y_{j \cdot k}^*(x_i^* - x_1^*),$$

it follows that

$$w_{i \cdot k}^* = (w_{i \cdot k} - w_{1 \cdot k}) - b_k c(x_i - x_1) + w_{1 \cdot k}^*.$$

And finally because  $w_{i \cdot k} + x_i y_{j \cdot k} + z_{j \cdot k} = w_{i \cdot k}^* + x_i^* y_{j \cdot k}^* + z_{j \cdot k}^*$ , it follows that

$$z_{j \cdot k}^* = z_{j \cdot k} + y_{j \cdot k}(x_1 - \frac{x_1^*}{c}) - b_k x_1^* + (w_{1 \cdot k} - w_{1 \cdot k}^*).$$

The bounds on  $b_k$  and  $c$  follow from the requirement that  $w_{i \cdot k}^*$ ,  $x_i^*$ ,  $y_{j \cdot k}^*$  and  $z_{j \cdot k}^*$  are between 0 and 1, for all  $i, j$  and  $k$ .  $\square$

*Remark 5.* For a  $C$ -class rate tree for two multiplicatively interacting factors, the theorem only needs to be modified for the bounds. To be specific, Inequalities (16) and (17) are not relevant if parameter  $y_{j \cdot k}$  is not bounded above by 1; the upper bound of Inequality (18) is not relevant if  $x_i$  is not bounded above by 1; Inequalities (12) and (20) are not relevant if parameter  $w_{i \cdot k}$  is not bounded above by 1; and Inequality (14) is not relevant if parameter  $z_{j \cdot k}$  is not bounded above by 1.

*Remark 6.* It is not possible to remove the interaction by a change of parameters. By assumption, with the old parameters the factors are effective interactively. That is, for some  $k$  there are levels  $(i, j)$  and  $(i^*, j^*)$  such that

$$(x_i - x_{i^*})(y_{j \cdot k} - y_{j^* \cdot k}) > 0.$$

A little algebra shows that if these parameters are transformed too in accordance with Theorem 2, the above interaction contrast is still positive.

*Remark 7.* The situation in which the factors are additive, that is, not effective interactively, is straightforward. Probabilities or frequencies are predicted by the tree through parameters in Equation (1). For a valid set of parameter values, the equation may have a nonzero multiplicative term, such as  $x_i y_k$  or  $x y_{j \cdot k}$ . The multiplicative term may depend on  $i$  or on  $j$ , but not on both. Such a term can be combined with the term  $w_{i \cdot k}$  if it depends on  $i$  or with  $z_{j \cdot k}$  if it depends on  $j$ .

*Remark 8.* One use of the previous theorem is to calculate the available degrees of freedom, which are a little complicated. With  $C$  classes,  $I$  levels of one factor and  $J$  levels of the other, there are  $CIJ$  observations. But for every  $i$  and every  $j$ , the sum over  $k$  of  $p_{ij \cdot k}$  is 1, so there are  $(C - 1)IJ$  independent observations.

To estimate, there is one parameter  $b$ , there are  $CI$  parameters  $w_{i \cdot k}$ ,  $I$  parameters  $x_i$ ,  $CJ$  parameters  $y_{j \cdot k}$  and  $CJ$  parameters  $z_{j \cdot k}$ , making  $(C+1)I+2CJ+1$  parameters to estimate. However, there are  $I + 2J$  equations relating these parameters. There are  $I$  equations of the form  $\sum_k (w_{i \cdot k} + x_i) = 1 - b$ ,  $J$  equations of the form  $\sum_k y_{j \cdot k} = 1$ , and  $J$  equations of the form  $\sum_k z_{j \cdot k} = b$ . The number of parameters to estimate is then  $CI - 2J(C - 1) + 1$ . But not all these need be estimated, because from Theorem 2, there are rescaling parameters which add to the degrees of freedom, namely  $c$ ;  $x_1^*$ ; for every  $k$ ,  $b_k$ ; and for every  $k$ ,  $w_{1 \cdot k}^*$ . There are  $2C+2$  rescaling parameters. However, the rescaling parameters are not independent because of the one constraint (21) in the Change of Parameter Bounds relating them. The number of independent rescaling parameters is then  $2C + 1$ .

The degrees of freedom is the number of independent observations minus the number of independent parameters to be fit, plus the number of independent rescaling pa-

rameters. That is, the degrees of freedom are

$$(C - 1)IJ - [CI - 2J(C - 1) + 1] + 2C + 1 = (I - 2)(JC - J - C).$$

The model cannot be tested with a small design. However, with three classes and three levels of each factor there are 3 degrees of freedom. With 4 classes and four levels of each factor there are 16 degrees of freedom.

### 2.2.3. *Uniqueness of structure.*

The preceding theorem is about the uniqueness of the parameters for a given tree structure. One can also ask about the uniqueness of the structure; the tree in Figure 5 is slightly different from the standard tree for multiplicative factors, but equivalent to it. What conditions are sufficient for such equivalence? Factors changing parameters in a processing or rate tree could interact in a complicated way, for example, as a sum of products of parameters. The simplest interaction has the form of a single product, so it is useful to know when this happens. The following two sections discuss uniqueness of structure. First, conditions are stated in Theorem 3 under which two process orders are possible in a standard tree for two multiplicatively interacting factors. In one tree parameters selectively influenced by Factor 1 precede parameters selectively influenced by Factor 2; in a different but equivalent tree the order is reversed. Theorem 4 gives conditions for an arbitrary two-indexed tree to be equivalent to the standard tree for two multiplicatively interacting factors; an important condition is that arcs associated with  $i$  and arcs associated with  $j$  are separated by a vertex.

#### 2.2.3.1. *Two possible orders for multiplicatively interacting factors.*

Suppose observations can be accounted for with a tree in which parameters changed by Factor 1 occur on a path preceding parameters changed by Factor 2. Sometimes an equivalent tree can be found in which parameters changed by the factors occur on a path in the opposite order, but sometimes no such tree is possible. The proof of Theorem 1 provides a procedure for finding parameters when parameters indexed by  $i$  precede parameters indexed by  $j$ . The procedure can be used analogously to find parameters when parameters indexed by  $j$  precede parameters indexed by  $i$ . The

following theorem shows the relations between the parameters necessary and sufficient for this to be possible. The conditions are given in two equivalent ways, each gives parameters for one order in terms of parameters for the reverse order. Bounds must be satisfied to keep parameters between 0 and 1. These are the Change of Parameter Order Bounds I and II, in Tables 2 and 3, respectively.

**Theorem 3.** *Suppose  $\{P_{(k)}\}_C$  is a probability matrix stack generated by two factors that are effective interactively. The following three statements are equivalent.*

(1) *Probability matrix stack  $\{P_{(k)}\}_C$  is produced by the two factors selectively influencing parameters in the standard tree for multiplicatively interacting factors with an arc associated with  $i$  preceding arcs associated with  $j$ , and  $\{P_{(k)}\}_C$  is also produced by the two factors selectively influencing parameters in the standard tree for multiplicatively interacting factors with an arc associated with  $j$  preceding arcs associated with  $i$ .*

(2) *For all  $k$ ,  $1 \leq k \leq C$ , there are probabilities  $w_{i \cdot k}$ ,  $w'_{i \cdot k}$ ,  $x_i$ ,  $x'_{i \cdot k}$  for all  $i$ ,  $1 \leq i \leq I$ , and probabilities  $z_{j \cdot k}$ ,  $z'_{j \cdot k}$ ,  $y_{j \cdot k}$ ,  $y'_j$  for all  $j$ ,  $1 \leq j \leq J$ , such that for all  $i$ ,  $j$  and  $k$ ,*

$$p_{ij \cdot k} = w_{i \cdot k} + x_i y_{j \cdot k} + z_{j \cdot k} = w'_{i \cdot k} + y'_j x'_{i \cdot k} + z'_{j \cdot k};$$

*with for every  $j$   $\sum_k y_{j \cdot k} = 1$  and  $\sum_k z_{j \cdot k} = b$ ,*

*and for every  $i$   $\sum_k x'_{i \cdot k} = 1$  and  $\sum_k w'_{i \cdot k} = b'$ .*

*There are constants  $c_k$  such that*

*for any  $k$  for which the factors are effective intractively, for every  $j$*

$$y'_j = \frac{y_{j \cdot k} - y_{1 \cdot k}}{c_k}$$

*and for every  $k$ ,  $i$  and  $j$*

$$x'_{i \cdot k} = c_k [x_i - x_{i^*(k)}]$$

$$w'_{i \cdot k} = w_{i \cdot k} - w_{1 \cdot k} + w'_{1 \cdot k} - c_k y'_{j^*(k)} x_i$$

$$z'_{j \cdot k} = z_{j \cdot k} + z'_{j^*(k) \cdot k} - z_{j^*(k) \cdot k} + x_{i^*(k)} y_{j \cdot k};$$

*where*

*1 is a value of  $i$  such that  $x_1 = 0$ ,*

1 is a value of  $j$  such that  $y'_1 = 0$ ,

and for all  $k$ ,

$i^*(k)$  is a value of  $i$  such that  $x'_{i \cdot k}$  takes on value 0 at  $i^*(k)$ ,

$j^*(k)$  is a value of  $j$  such that  $y_{j \cdot k}$  takes on value 0 at  $j^*(k)$ ;

further, Change of Order Bounds I are satisfied.

(3) Probability matrix stack  $\{P_{(k)}\}_C$  is produced by two factors selectively influencing parameters in a standard tree for multiplicatively interacting factors with an arc associated with  $i$  preceding arcs associated with  $j$  in which

$$p_{ij \cdot k} = w_{i \cdot k} + x_i y_{j \cdot k} + z_{j \cdot k},$$

for probabilities  $w_{i \cdot k}$ ,  $x_i$ ,  $y_{j \cdot k}$  and  $z_{j \cdot k}$ , with  $\sum_k y_{j \cdot k} = 1$  and  $\sum_k z_{j \cdot k} = b$ .

There are constants  $c_k$  such that

for any  $k$  for which the factors are effective interactively, for every  $i$ ,  $x_i$  has the form

$$x_i = \frac{x(i, k)}{c_k} + h(k),$$

and for every  $k$  for every  $j$ ,  $y_{j \cdot k}$  has the form

$$y_{j \cdot k} = c_k[y(j) - g(k)];$$

further, Change of Order Parameter Bounds II are satisfied.

*Proof.* (a) To see that Statement (1) implies Statement (2), suppose  $\{P_{(k)}\}_C$  is produced by two factors selectively influencing parameters in a standard tree for multiplicatively interacting factors with an arc indexed by  $i$  preceding arcs indexed by  $j$ , and is also produced by two factors selectively influencing parameters in a standard tree for multiplicatively interacting factors with an arc indexed by  $j$  preceding arcs indexed by  $i$ .

According to Theorem 1, probabilities  $w_{i \cdot k}$ ,  $z_{j \cdot k}$ ,  $x_i$  and  $y_{j \cdot k}$  can be assigned such that  $p_{ij \cdot k} = w_{i \cdot k} + x_i y_{j \cdot k} + z_{j \cdot k}$ . From Remarks 3 and 4 following Theorem 1, parameter values for  $x_i$  can be chosen so  $x_1 = 0$ , and for every  $k$ , values of  $y_{j \cdot k}$  can be chosen so for some  $j^*(k)$ ,  $y_{j^*(k) \cdot k} = 0$ .

Also, with the same procedure, for the tree in which the arc indexed by  $j$  precedes the arcs indexed by  $i$  we can assign probabilities  $w'_{i \cdot k}$ ,  $z'_{j \cdot k}$ ,  $x'_{i \cdot k}$  and  $y'_j$  such that

$p_{ij \cdot k} = w'_{i \cdot k} + y'_j x'_{i \cdot k} + z'_{j \cdot k}$ . Note that use of Remarks 3 and 4 to set  $x_1 = 0$  is based on establishing an order on the levels  $i$ . This does not establish an order on the levels  $j$ . Then parameter values can be chosen so  $y'_1 = 0$  and for every  $k$  there is a level  $i^*(k)$ , such that  $x'_{i^*(k) \cdot k} = 0$ .

Then for any  $i$  and  $j$ , for every class  $k$ ,

$$\begin{aligned} & p_{ij \cdot k} - p_{i^*(k)j \cdot k} - p_{ij^*(k) \cdot k} + p_{i^*(k)j^*(k) \cdot k} \\ &= (x_i - x_{i^*(k)})(y_{j \cdot k} - y_{j^*(k) \cdot k}) \\ &= (x_i - x_{i^*(k)})y_{j \cdot k}. \end{aligned}$$

Also, for this same interaction contrast

$$\begin{aligned} & p_{ij \cdot k} - p_{i^*(k)j \cdot k} - p_{ij^*(k) \cdot k} + p_{i^*(k)j^*(k) \cdot k} \\ &= (y'_j - y'_{j^*(k)})(x'_{i \cdot k} - x'_{i^*(k) \cdot k}) \\ &= (y'_j - y'_{j^*(k)})x'_{i \cdot k}. \end{aligned}$$

Thus,

$$(x_i - x_{i^*(k)})y_{j \cdot k} = (y'_j - y'_{j^*(k)})x'_{i \cdot k}. \quad (7)$$

Consider any  $k$ . Choose  $i$  such that  $x_i \neq x_{i^*(k)}$ . Such an  $i$  exists, otherwise, the value of  $x_i$  is the same for every  $i$  and the factors are not effective interactively for any  $k$ . Likewise, choose  $j$  such that  $y'_j \neq y'_{j^*(k)}$ . For any  $i$  and  $j$  so chosen, from Equation (7),

$$\frac{x'_{i \cdot k}}{x_i - x_{i^*(k)}} = \frac{y_{j \cdot k}}{y'_j - y'_{j^*(k)}} = c_k. \quad (8)$$

Because the ratios above do not depend on  $i$  or  $j$ , they equal a constant  $c_k$  that depends only on  $k$ .

We have immediately for  $i$  such that  $x_i \neq x_{i^*(k)}$ ,

$$x'_{i \cdot k} = c_k(x_i - x_{i^*(k)}).$$

Now consider  $i$  such that  $x_i = x_{i^*(k)}$  but  $j$  is as before such that  $y'_j \neq y'_{j^*(k)}$ . The left hand side of Equation (7) is 0; for the right side to be 0,  $x'_{i \cdot k} = 0$ . Therefore, the equation above also gives the correct value of  $x'_{i \cdot k}$  for  $i$  such that  $x_i = x_{i^*(k)}$ , and hence for all  $i$ .



To derive an expression for  $y'_j$ , consider a class  $k$  for which the factors are effective interactively. From Equation (8), if  $c_k = 0$ , then  $y_{j \cdot k} = 0$  for all  $j$  and the factors are not effective interactively for  $k$ . Hence  $c_k \neq 0$  and it follows from Equation (8) that for every  $j$

$$y'_j = y_{j \cdot k} / c_k + y'_{j^*(k)}.$$

For  $j = 1$ ,

$$y'_1 = 0 = \frac{y_{1 \cdot k}}{c_k} + y'_{j^*(k)}.$$

Hence,  $y'_{j^*(k)} = -y_{1 \cdot k} / c_k$ , so

$$y'_j = \frac{y_{j \cdot k} - y_{1 \cdot k}}{c_k}.$$

We now derive expressions for  $w'_{i \cdot k}$  and  $z'_{j \cdot k}$ . It follows from the derivations above that

$$x'_{i \cdot k} y'_j = c_k x_i y'_{j^*(k)} + x_i y_{j \cdot k} - c_k y'_{j^*(k)} x_{i^*(k)} - x_{i^*(k)} y_{j \cdot k}.$$

So

$$w'_{i \cdot k} + z'_{j \cdot k} = w_{i \cdot k} + z_{j \cdot k} - c_k x_i y'_{j^*(k)} + c_k y'_{j^*(k)} x_{i^*(k)} + x_{i^*(k)} y_{j \cdot k}.$$

Because  $w'_{i \cdot k}$  does not depend on  $j$  and  $z'_{j \cdot k}$  does not depend on  $i$ , there exist numbers  $a_k$  and  $b_k$  such that

$$w'_{i \cdot k} - w_{i \cdot k} + c_k x_i y'_{j^*(k)} = a_k$$

and

$$z'_{j \cdot k} - z_{j \cdot k} - x_{i^*(k)} y_{j \cdot k} = b_k,$$

with

$$a_k + b_k = c_k y'_{j^*(k)} x_{i^*(k)}.$$

In particular, when  $i = 1$ ,  $x_1 = 0$ , thus

$$a_k = w'_{1 \cdot k} - w_{1 \cdot k}.$$

And so  $w'_{i \cdot k} = w_{i \cdot k} + w'_{1 \cdot k} - w_{1 \cdot k} - c_k x_i y'_{j^*(k)}$ .

Similarly, when  $j = j^*(k)$ ,  $y_{j^*(k) \cdot k} = 0$ , thus

$$b_k = z'_{j^*(k) \cdot k} - z_{j^*(k) \cdot k}.$$

And so  $z'_{j \cdot k} = z_{j \cdot k} + z'_{j^*(k) \cdot k} - z_{j^*(k) \cdot k} + x_{i^*(k)} y_{j \cdot k}$ .

For the bounds, Equation (22) follows from

$$\begin{aligned} & w'_{i \cdot k} + y'_j x'_{i \cdot k} + z'_{j \cdot k} \\ = & w_{i \cdot k} + x_i y_{j \cdot k} + z_{j \cdot k} + w'_{1 \cdot k} - w_{1 \cdot k} + z'_{j^*(k) \cdot k} - z_{j^*(k) \cdot k} - c_k x_{i^*(k)} y'_{j^*(k)}. \end{aligned}$$

Therefore,

$$w'_{1 \cdot k} - w_{1 \cdot k} + z'_{j^*(k) \cdot k} - z_{j^*(k) \cdot k} - c_k x_{i^*(k)} y'_{j^*(k)} = 0.$$

The other bounds on  $c_k$  follow after some algebra from the requirements that  $w'_{i \cdot k}$ ,  $x'_{i \cdot k}$ ,  $z'_{j \cdot k}$  and  $y'_j$  are probabilities. Note that since  $\sum_{k=1}^C c_k = 0$ ,  $c_k$  is negative for some  $k$ 's.

(b) We now show that Statement (2) implies Statement (3). Suppose Statement (2) is true. Then for any  $k$

$$x'_{i \cdot k} = c_k (x_i - x_{i^*(k)}).$$

Choose  $k$  for which the factors are effective interactively. Then  $c_k \neq 0$ . Then

$$x_i = \frac{x(i, k)}{c_k} + x_{i^*(k)},$$

so  $x_i$  has the form

$$x_i = \frac{x(i, k)}{c_k} + h(k).$$

Now consider  $y_{j \cdot k}$ . From Statement (2) for any  $k$  for which the factors are effective interactively

$$y'_j = \frac{y_{j \cdot k}}{c_k} - \frac{y_{1 \cdot k}}{c_k}.$$

Then

$$y'_{j^*(k)} = \frac{y_{j^*(k) \cdot k}}{c_k} - \frac{y_{1 \cdot k}}{c_k} = -\frac{y_{1 \cdot k}}{c_k},$$

so

$$y'_j = \frac{y_{j \cdot k}}{c_k} + y'_{j^*(k)}.$$

For such a  $k$

$$y_{j \cdot k} = c_k(y'_j - y'_{j^*(k)}). \quad (9)$$

We now show the above equation is true for  $k$  for which the factors are not effective interactively, and thus is true for all  $k$ . Consider  $k$  for which the factors are not effective interactively. For every  $i$  and  $j$ ,

$$(x_i - x_{i^*(k)})(y_{j \cdot k} - y_{j^*(k) \cdot k}) = 0.$$

It is not possible that for every  $i$

$$x_i - x_{i^*(k)} = 0,$$

if so the factors would not be effective interactively for any  $k$ . It must be that for every  $j$

$$y_{j \cdot k} - y_{j^*(k) \cdot k} = 0.$$

Because  $j^*(k)$  is a value of  $j$  such that  $y_{j^*(k) \cdot k} = 0$ , it follows that

$$y_{j \cdot k} = 0,$$

for every  $j$  when the factors are not effective interactively for  $k$ .

Further, for  $k$  for which the factors are not effective interactively, for all  $i$  and  $j$

$$(x'_{i \cdot k} - x_{i^*(k) \cdot k})(y'_j - y'_{j^*(k)}) = 0.$$

Because  $x_{i^*(k) \cdot k} = 0$ ,

$$x'_{i \cdot k}(y'_j - y'_{j^*(k)}) = 0.$$

It is not possible that for all  $j$ ,  $y'_j - y'_{j^*(k)} = 0$  because then the factors would not be effective interactively for any  $k$ . It must be that  $x'_{i \cdot k} = 0$  for every  $i$  when the factors are not effective interactive interactively at  $k$ . From Statement (2), for any  $k$ ,

$$x'_{i \cdot k} = c_k(x_i - x_{i^*(k)}).$$

Consider  $k$  for which the factors are not effective interactively. For the left hand side to be 0 for every  $i$ , it must be that  $c_k = 0$ . Then Equation (9) for  $y_{j \cdot k}$  gives the correct value of 0 for  $k$ .

Therefore, for every  $k$ ,  $y_{j \cdot k}$  can be rewritten in the following form:

$$y_{j \cdot k} = c_k[y(j) - g(k)].$$

Further,

$$0 \leq x(i, k), h(k), y(j), g(k) \leq 1.$$

A little algebra shows Change of Order Bounds II follow. In particular, for any  $j$ ,  $\sum_{k=1}^C y_{j \cdot k} = 1$ . So  $1 = y(j) \sum_{k=1}^C c_k - \sum_{k=1}^C c_k g(k)$ .

Because  $\sum_{k=1}^C c_k g(k)$  and  $\sum_{k=1}^C c_k$  are both independent of  $j$ ,  $\sum_{k=1}^C c_k = 0$ . Otherwise,  $y(j)$  is a constant and the factors are not effective interactively in the standard tree for multiplicatively interacting factors with an arc indexed by  $j$  preceding arcs indexed by  $i$ . Thus,  $\sum_{k=1}^C c_k = 0$ ,  $\sum_{k=1}^C c_k g(k) = -1$ .

Also, for any  $i$ ,  $\sum_{k=1}^C x'_{i \cdot k} = 1$ . So  $1 = x_i \sum_{k=1}^C c_k - \sum_{k=1}^C c_k h(k)$ . Similarly,  $\sum_{k=1}^C c_k h(k) = -1$ .

(c) Finally to show that Statement (3) implies Statement (1), we have that  $\{P_{(k)}\}_C$  is produced by two factors effective interactively, selectively influencing parameters in a standard tree for multiplicatively interacting factors with an arc indexed by  $i$  preceding arcs indexed by  $j$  in which  $p_{i \cdot j \cdot k} = w_{i \cdot k} + x_i y_{j \cdot k} + z_{j \cdot k}$ .

As  $x_i = \frac{x(i, k)}{c_k} + h(k)$  and  $y_{j \cdot k} = c_k[y(j) - g(k)]$ ,  $0 \leq x(i, k), y(j) \leq 1$ , let probabilities

$$\begin{aligned} x'_{i \cdot k} &= x(i, k), \\ y'_j &= y(j). \end{aligned}$$

Let  $j^{**}$  be chosen such that  $y(j)$  takes on its smallest value at  $j^{**}$ , and for every  $k$ , let  $i^{**}(k)$  be chosen so that  $x(i, k)$  takes on its smallest value at  $i^{**}(k)$ .

Then the interaction for any combination of levels  $(i, j)$  with respect to the level combination  $(i^{**}(k), j^{**})$  is

$$\begin{aligned} p_{i \cdot j \cdot k} - p_{i \cdot j^{**} \cdot k} &= p_{i^{**}(k) \cdot j \cdot k} - p_{i^{**}(k) \cdot j^{**} \cdot k} \\ &= [x_i - x_{i^{**}(k)}](y_{j \cdot k} - y_{j^{**} \cdot k}) \\ &= \left[ \frac{x(i, k)}{c_k} - \frac{x(i^{**}(k), k)}{c_k} \right] [c_k y(j) - c_k y(j^{**})] \\ &= [x(i, k) - x(i^{**}(k), k)][y(j) - y(j^{**})] \\ &= r_j s_{i \cdot k}, \end{aligned}$$

for nonnegative  $r_j$  and  $s_{i \cdot k}$ .

Further, from Change of Parameter Order Bounds II, Inequality (40),

$$p_{ij \cdot k} \geq x(i, k)y(j).$$

Hence,

$$p_{ij \cdot k} \geq [x(i, k) - x(i^{**}(k), k)][y(j) - y(j^{**})].$$

And

$$\sum_{k=1}^C x'_{i \cdot k} = \sum_{k=1}^C x(i, k) = x_i \sum_{k=1}^C c_k - \sum_{k=1}^C c_k h(k) = 1.$$

And thus

$$\sum_{k=1}^C s_{i \cdot k} = 1 - \sum_{k=1}^C x(i^{**}(k), k) = a,$$

where  $a$  is a constant and  $0 < a \leq 1$ . Otherwise, the factors are not effective interactively for any class  $k$ .

Then, according to Theorem 1, probability matrix stack  $\{P_{(k)}\}_C$  is also produced by a standard tree for multiplicatively interacting factors with an arc indexed by  $j$  preceding arcs indexed by  $i$ .  $\square$

*Remark 9.* The conditions under which two process orders are possible in a  $C$ -class rate tree for two multiplicatively interacting factors can be established analogously. Statements need to be modified using appropriate terms (for example,  $\{P_{(k)}\}_C$  is now a rate matrix stack), but the major modifications are on the parameter bounds. No parameter needs to be bounded above by 1 in a rate model. Modifications are summarized in Table 4.

*Remark 10.* The preceding theorem is not relevant when the factors are not effective interactively, of course.

*Example 4.* Suppose two factors selectively influence parameters in the standard tree for multiplicative factors. If an arc indexed by  $i$  precedes arcs indexed by  $j$ , it must be that  $\sum_k w'_{i \cdot k}$  is constant. Consider the simple model in Figure 8. Suppose there were only two response classes, with response 1 in place of response 3. It is easy to check by redrawing the tree that the order of parameters indexed by  $i$  and those indexed by

$j$  can be reversed. But with three classes reversal is not possible. For the tree with an arc indexed by  $j$  preceding arcs indexed by  $i$ ,  $w'_{i.1} = 1 - x_i$ ,  $w'_{i.2} = w'_{i.3} = 0$ . Then  $\sum_k w'_{i.k}$  depends on  $i$  and is not constant.

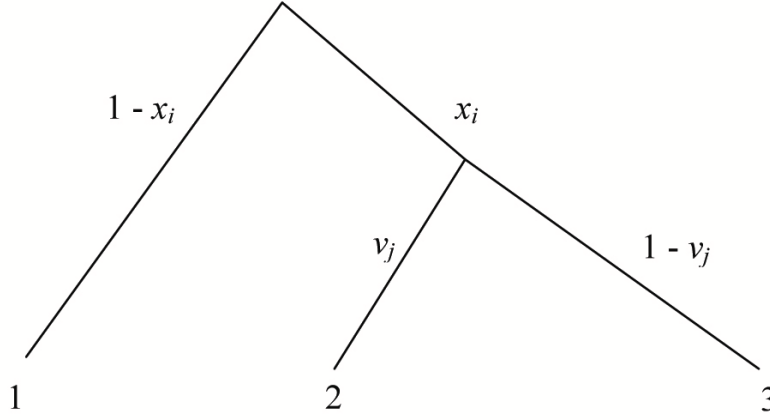


Figure 8: A simple processing tree model

#### 2.2.3.2. Arbitrary trees, equivalent to the standard tree.

The following theorem shows that if two multiplicatively interacting factors influence processes in an arbitrary  $C$ -class processing tree with certain properties, the tree is equivalent to a standard  $C$ -class processing tree for two multiplicatively interacting factors.

Before stating the theorem, we explain a special type of vertex. In a processing or rate tree for two multiplicatively interacting factors, there are three types of paths from the root to a terminal vertex associated with some response. In the first type there are arcs whose parameters either do not change or change only with Factor 1 (Type I path); in the second type there are arcs whose parameters either do not change or change only with Factor 2 (Type II path); and in the third type, which makes the two factors interact, there are arcs whose parameters change with Factor 1, arcs whose parameters change with Factor 2, and possibly arcs whose parameters not change (Type III path).

In a Type III path, it may be that the arcs are ordered in such a way that all the arcs

indexed by  $j$  precede the arcs indexed by  $i$ , or that all the arcs indexed by  $i$  precede the arcs indexed by  $j$ . If it happens that arcs along the path can be ordered in this way, we can always find at least one vertex that separates the arcs indexed by  $i$  and those indexed by  $j$ . Such a vertex is called a *separating vertex* and denoted  $o$ . In other words, arcs preceding the separating vertex are only indexed by  $i$  (or only indexed by  $j$ ) or are constant parameters while the arcs following the separating vertex are only indexed by  $j$  (or only indexed by  $i$ ) or are constant parameters.

Given an arc  $a$  and a vertex  $v$ , if there is no path from the root to a terminal vertex that contains both  $a$  and  $v$  we say  $a$  and  $v$  are *unordered*.

**Theorem 4.** *Suppose  $\{P_{(k)}\}_C$  is a probability matrix stack produced by factors selectively influencing parameters in an arbitrary two-indexed processing tree with  $C$  classes. Suppose the factors are effective interactively. Suppose the following two conditions are satisfied:*

- (1) *All arcs indexed by  $i$  precede all arcs indexed by  $j$  along all Type III paths;*
- (2) *There are distinct separating vertices  $o_1, \dots, o_N$  such that each Type III path contains exactly one of them and for each separating vertex  $o_n$ , there is a number  $\alpha_n$  such that the probability of starting from the root and reaching  $o_n$  can be written  $\alpha_n \delta(i)$ , where  $\delta(i)$  is a function of  $i$ .*

*Then  $\{P_{(k)}\}_C$  is produced by an equivalent standard  $C$ -class processing tree for multiplicatively interacting factors, with an arc indexed by  $i$  preceding arcs indexed by  $j$ .*

*Proof.* Clearly, in an arbitrary two-indexed processing tree as described in Condition (1), every arc indexed by  $i$  either is unordered with respect to every separating vertex  $o_n$  or precedes one or more of them. And, every arc indexed by  $j$  either is unordered with respect to every separating vertex  $o_n$  or follows one or more of them.

There is a path from the root to each separating vertex  $o_n$ . On each path, some arcs are indexed by  $i$  and no arcs are indexed by  $j$ . According to Condition (2), the probability of the path from the root to the vertex  $o_n$  is  $\alpha_n \delta(i)$ .

Consider the single subtree consisting of the vertex  $o_n$  as the root and all its descendants. For a given response class  $k$ , suppose there are  $T_{(k,n)}$  terminal vertices in the

subtree, with each terminal vertex denoted as  $t_{(m;k,n)}$ , where  $m = 1, \dots, T_{(k,n)}$ . For each terminal vertex  $m$ , there is a path from  $o_n$  to  $t_{(m;k,n)}$ . The probability of the path may depend on  $j$  but not  $i$ . Let the probability be  $f_n(j, t_{(n,m,k)})$ . Thus in the subtree, the probability of a response  $R_k$  is

$$\sum_{m=1}^{T_{(k,n)}} f_n(j, t_{(n,m,k)}).$$

Note that this probability depends on  $n$ ,  $k$  and  $j$ .

There are four ways to start from the root of the original tree and reach a terminal vertex for response class  $k$ .

(a) A path may start from the root and reach a terminal vertex for class  $k$ , with no arcs on the path indexed by  $i$  or  $j$ . Suppose there are  $Q_k$  such paths and let the probability of path  $q_k$  be  $p_{(k,q_k)}$ . Then, the probability of reaching a terminal vertex for class  $k$  by such a path is

$$\sum_{q_k=1}^{Q_k} p_{(k,q_k)}.$$

(b) A path may start from the root and reach a terminal vertex for class  $k$ , with some of the arcs on the path indexed by  $i$  but no arcs indexed by  $j$ . Suppose there are  $U_k$  such paths, and let the probability of path  $u_k$  be  $g(i, u_k)$ . Then the probability of reaching a terminal vertex for class  $k$  by such a path is

$$\sum_{u_k=1}^{U_k} g(i, u_k).$$

(c) A path may start from the root and reach a terminal vertex for class  $k$ , on which some of the arcs are indexed by  $j$  but none are indexed by  $i$ . Suppose there are  $V_k$  such paths for class  $k$ , and let the probability of path  $v_k$  be  $h(j, v_k)$ . Then the probability of reaching a terminal vertex for class  $k$  by such paths is

$$\sum_{v_k=1}^{V_k} h(j, v_k).$$

(d) A path may start from the root and reach a terminal vertex for class  $k$ , with some of the arcs on the path indexed by  $i$  and some indexed by  $j$ . Such a path is one



that starts from the root, reaches one of the separating vertices  $o_n$ ,  $n = 1, \dots, N$ , and finally reaches a terminal vertex for response class  $k$ .

From earlier statements, the probability of reaching a terminal vertex for class  $k$  by such a path is

$$\alpha_n \delta(i) \sum_{m=1}^{T_{(k,n)}} f_n(j, t_{(n,m,k)}).$$

Over all such paths, the probability of reaching a terminal vertex for class  $k$  is

$$\sum_{n=1}^N [\alpha_n \delta(i) \sum_{m=1}^{T_{(k,n)}} f_n(j, t_{(n,m,k)})].$$

The four ways of making a response  $k$  are mutually exclusive. Hence, the probability of a response  $k$  when Factor 1 is at level  $i$  and Factor 2 is at level  $j$  is

$$\begin{aligned} p_{ij \cdot k} &= \sum_{q_k=1}^{Q_k} p_{(k,q_k)} + \sum_{u_k=1}^{U_k} g(i, u_k) + \sum_{v_k=1}^{V_k} h(j, v_k) + \sum_{n=1}^N [\alpha_n \delta(i) \sum_{m=1}^{T_{(k,n)}} f_n(j, t_{(n,m,k)})] \\ &= \sum_{q_k=1}^{Q_k} p_{(k,q_k)} + \sum_{u_k=1}^{U_k} g(i, u_k) + \sum_{v_k=1}^{V_k} h(j, v_k) + \delta(i) \left\{ \sum_{n=1}^N [\alpha_n \sum_{m=1}^{T_{(k,n)}} f_n(j, t_{(n,m,k)})] \right\}. \end{aligned}$$

For all  $i$ , let

$$x_i = \delta(i) \left( \sum_{n=1}^N \alpha_n \right),$$

$$w_{i \cdot k} = \sum_{q_k=1}^{Q_k} p_{(k,q_k)} + \sum_{u_k=1}^{U_k} g(i, u_k),$$

and for all  $j$ , let

$$y_{j \cdot k} = \frac{\sum_{n=1}^N [\alpha_n \sum_{m=1}^{T_{(k,n)}} f_n(j, t_{(n,m,k)})]}{\sum_{n=1}^N \alpha_n},$$

$$z_{j \cdot k} = \sum_{v_k=1}^{V_k} h(j, v_k).$$

Then clearly,  $p_{ij \cdot k} = w_{i \cdot k} + x_i y_{j \cdot k} + z_{j \cdot k}$ .

For any of the  $N$  subtrees, the sum of the probabilities for all response classes when Factor 2 is at any particular level  $j$  is 1, i.e.

$$\sum_{k=1}^C \sum_{m=1}^{T(k,n)} f_n(j, t_{(n,m,k)}) = 1.$$

Thus,

$$\begin{aligned} \sum_{k=1}^C y_{j \cdot k} &= \sum_{k=1}^C \left\{ \frac{\sum_{n=1}^N [\alpha_n \sum_{m=1}^{T(k,n)} f_n(j, t_{(n,m,k)})]}{\sum_{n=1}^N \alpha_n} \right\} \\ &= \frac{\sum_{n=1}^N [\alpha_n \sum_{k=1}^C \sum_{m=1}^{T(k,n)} f_n(j, t_{(n,m,k)})]}{\sum_{n=1}^N \alpha_n} \\ &= \frac{\sum_{n=1}^N \alpha_n}{\sum_{n=1}^N \alpha_n} = 1. \end{aligned}$$

Also, because the sum of the probabilities of all the children of a vertex is 1,

$$x_i = \delta(i) \sum_{n=1}^N \alpha_n = \sum_{n=1}^N \delta(i) \alpha_n \leq 1.$$

Thus  $0 \leq x_i, y_{j \cdot k} \leq 1$ .

Because  $p_{i \cdot k} = w_{i \cdot k} + x_i y_{j \cdot k} + z_{j \cdot k}$ , then  $0 \leq w_{i \cdot k} + z_{j \cdot k} \leq 1$ .

Because both  $w_{i \cdot k}$  and  $z_{j \cdot k}$  are non-negative in the above inequality, they are both bounded above by 1. Hence they are both probabilities, i.e.,

$$0 \leq w_{i \cdot k}, z_{j \cdot k} \leq 1.$$

Therefore, an arbitrary two-indexed processing tree as described in Theorem 4 corresponds to an equivalent standard  $C$ -class processing tree for two multiplicatively interacting factors, with an arc indexed by  $i$  preceding arcs indexed by  $j$ .  $\square$

*Remark 11.* Now consider a  $C$ -class rate tree for two multiplicatively interacting factors. The three types of paths and separating vertices can be defined in the same way, and Theorem 4 holds with only minor modifications to eliminate the requirement that each parameter is bounded above by 1.

*Remark 12.* If the factors are not effective interactively, there are no Type III paths, and Theorem 4 applies with straightforward modifications.

*Example 5. Weapon Misidentification*

In the Quad Model of Figure 6, suppose parameter  $A$  is indexed by  $i$  and parameter  $G$  is indexed by  $j$ . On the left side of the tree, the ending vertex of the arc labeled  $1 - A_i$  separates arcs indexed by  $i$  from arcs indexed by  $j$ . The path from the root to this vertex has probability  $(1 - OB)(1 - A_i)$ . On the right side of the tree, the ending vertex of the arc labeled  $1 - A_i$  also separates arcs indexed by  $i$  from arcs indexed by  $j$ . The probability of this path is  $OB * (1 - C)(1 - A_i)$ . This tree meets the conditions of Theorem 4, and therefore is equivalent to the standard tree for multiplicatively interacting factors.

As mentioned in the introduction, the Quad Model can be considered a mixture of two models, and it is worth considering when a mixture of models is equivalent to a single standard tree with multiplicatively interacting factors.

*2.2.4. Mixtures of trees.*

Consider a set of processing or rate trees. Suppose when Factor 1 is at level  $i$  and Factor 2 is at level  $j$ , tree  $t$  predicts the probability or rate for class  $k$  as  $p_{tij \cdot k}$ . Consider a mixture of the trees, in which tree  $t$  is selected with probability  $\tau_t$ . We say the probability (or rate) matrix stack  $\{P_{(k)}\}_C$  is produced by the mixture of trees if

$$p_{ij \cdot k} = \sum_t \tau_t p_{tij \cdot k}.$$

The following theorem shows that if two multiplicatively interacting factors selectively influence processes in every individual two-indexed tree in a mixture of trees, the mixture is equivalent to a standard  $C$ -class processing tree for two multiplicatively interacting factors, if certain conditions hold. In particular, Statement (2) of Theorem 4 is revised so it applies with the same value of  $\delta(i)$  for each tree of the mixture.

**Theorem 5.** *Suppose probability matrix stack  $\{P_{(k)}\}_C$  is produced by a mixture of  $C$ -class processing trees, each tree with some arcs indexed by  $i$ , some by  $j$  and no arc indexed by both  $i$  and  $j$ . Suppose the following conditions hold.*

- (1) *All arcs indexed by  $i$  precede all arcs indexed by  $j$  along all Type III paths in each tree.*
- (2) *For each tree  $t$  there are distinct separating vertices  $o_1, \dots, o_{N_t}$  such that each Type*

III path contains exactly one of them and for each separating vertex  $o_n$ , there is a number  $\alpha_{tn}$  such that the probability of starting from the root and reaching  $o_n$  can be written  $\alpha_{tn}\delta(i)$ , where  $\delta(i)$  is a function of  $i$ .

(3) For each tree, the factors are effective interactively.

Then  $\{P_{(k)}\}_C$  is produced by an equivalent standard tree for two multiplicatively interacting factors, with an arc associated with  $i$  preceding arcs associated with  $j$ .

*Proof.* Let  $p_{tij \cdot k}$  be the probability of response class  $R_k$  for tree  $t$ , when Factor 1 is at level  $i$  and Factor 2 is at level  $j$ . According to Theorem 4, for each tree  $t$  there exists an equivalent standard tree for multiplicatively interacting factors, such that

$$p_{tij \cdot k} = w_{ti \cdot k} + x_{ti}y_{tj \cdot k} + z_{tj \cdot k}.$$

From the proof of Theorem 4,  $x_{ti}$  in the standard tree for tree  $t$  can be set equal to

$$\delta(i) \sum_{n=1}^{N_t} \alpha_{tn}.$$

So let  $x_{ti} = c_t \delta(i)$ , where

$$c_t = \sum_{n=1}^{N_t} \alpha_{tn}.$$

Suppose tree  $t$  is selected with probability  $\tau_t$ .

Over the mixture of trees,

$$\begin{aligned} p_{ij \cdot k} &= \sum_t \tau_t p_{tij \cdot k} \\ &= \sum_t \tau_t w_{ti \cdot k} + \sum_t \tau_t x_{ti} y_{tj \cdot k} + \sum_t \tau_t z_{tj \cdot k} \\ &= \sum_t \tau_t w_{ti \cdot k} + \delta(i) \sum_t c_t \tau_t y_{tj \cdot k} + \sum_t \tau_t z_{tj \cdot k}. \end{aligned}$$

Let  $u_{j \cdot k} = \sum_t c_t \tau_t y_{tj \cdot k}$ .

Then  $p_{ij \cdot k} = \sum_t \tau_t w_{ti \cdot k} + \delta(i) u_{j \cdot k} + \sum_t \tau_t z_{tj \cdot k}$ .

For every  $k$ , let  $j^*(k)$  be chosen so  $u_{j \cdot k}$  takes on its smallest value at  $j^*(k)$ .

Also, let  $i^*$  be chosen so  $\delta(i)$  takes on its smallest value at  $i^*$ .

Then for every  $i, j$ , and  $k$ ,

$$\begin{aligned}
p_{ij \cdot k} - p_{ij^*(k) \cdot k} - p_{i^* j \cdot k} &+ p_{i^* j^*(k) \cdot k} \\
&= \delta(i)u_{j \cdot k} - \delta(i)u_{j^*(k) \cdot k} - \delta(i^*)u_{j \cdot k} + \delta(i^*)u_{j^*(k) \cdot k} \\
&= [\delta(i) - \delta(i^*)](u_{j \cdot k} - u_{j^*(k) \cdot k}).
\end{aligned}$$

Clearly,  $p_{ij \cdot k} \geq \delta(i)u_{j \cdot k} \geq [\delta(i) - \delta(i^*)](u_{j \cdot k} - u_{j^*(k) \cdot k})$ .

Let  $r_i = \delta(i) - \delta(i^*)$  and  $s_{j \cdot k} = u_{j \cdot k} - u_{j^*(k) \cdot k}$ , then

$$p_{ij \cdot k} - p_{ij^*(k) \cdot k} - p_{i^* j \cdot k} + p_{i^* j^*(k) \cdot k} = r_i s_{j \cdot k}.$$

Further,

$$\begin{aligned}
\sum_{k=1}^C s_{j \cdot k} &= \sum_k (u_{j \cdot k} - u_{j^*(k) \cdot k}) \\
&= \sum_k \sum_t \tau_t c_t y_{tj \cdot k} - \sum_k \sum_t \tau_t c_t y_{tj^*(k) \cdot k} \\
&= \sum_t \sum_k \tau_t c_t y_{tj \cdot k} - \sum_t \sum_k \tau_t c_t y_{tj^*(k) \cdot k} \\
&= \sum_t \tau_t c_t \sum_k y_{tj \cdot k} - \sum_t \tau_t c_t \sum_k y_{tj^*(k) \cdot k} \\
&= \sum_t \tau_t c_t - \sum_t \tau_t c_t \sum_k y_{tj^*(k) \cdot k}.
\end{aligned}$$

Note that  $\sum_{k=1}^C s_{j \cdot k}$  is not dependent on  $j$ . Let  $\sum_{k=1}^C s_{j \cdot k}$  be denoted as  $a$ , for all  $j$ .

Because  $u_{j^*(k) \cdot k} \leq u_{j \cdot k}$  for all  $j$  and  $k$ ,  $a \geq 0$ .

And as for at least one  $k$ ,  $u_{j^*(k) \cdot k}$  is not equal to  $u_{j \cdot k}$  (otherwise the factors are not effective interactively),  $a > 0$ .

Also, because  $a \leq \sum_{k=1}^C u_{j \cdot k} \leq 1$ , then  $0 < a \leq 1$ .

Therefore, the conditions of Theorem 1 are satisfied and  $\{P_{(k)}\}_C$  is produced by a standard tree for two multiplicatively interacting factors with an arc indexed by  $i$  preceding arcs associated with  $j$ .  $\square$

*Remark 13.* A useful way in which Theorem 5 is general is that each tree in the mixture of trees may only allow some of the responses to be made; that is, each tree may have

terminal vertices associated with responses from only a proper subset of the response classes. Theorem 5 still holds, if we simply consider a tree in which there are no terminal vertices associated with a missing class as having probability 0 of producing that class of response. Also, it is easy to see that the theorem applies to rate trees with simple modifications.

*Remark 14.* It is also easy to see that if the factors are not effective interactively for any tree in the mixture of trees, the theorem applies with simple modifications.

### 3. Conclusion

Schweickert and Chen (2008) showed that if each of two factors change parameter values at exactly one vertex, only two trees need to be considered, the standard trees for unordered and ordered processes. Theorems on representation, choice of parameter values and uniqueness of structure were developed. The main purpose of this paper is to generalize the trees discussed in Schweickert and Chen (2008) to allow each factor to have effects at more than one vertex, to allow more than two response classes, and to allow parameters such as rates, that are not bounded above by 1. The generalizations complicate matters, but are manageable if certain conditions are met.

Trees of interest here can be considered as a combination of the standard unordered and ordered processing trees proposed by Schweickert and Chen (2008). It is straightforward to generalize the two-class standard tree for unordered processes of Schweickert and Chen (2008) to a tree with more than two classes. A standard  $C$ -class processing tree for unordered processes is illustrated in Figure 9. When Factor 1 is at level  $i$  and Factor 2 is at level  $j$ , the probability of the  $k$ th response class  $R_k$  is

$$p_{ij \cdot k} = \alpha w_{i \cdot k} + (1 - \alpha) z_{j \cdot k}.$$

In this tree, the two factors are not multiplicatively interacting.

In the standard tree for ordered processes of Schweickert and Chen (2008), generalized to more than two classes, the probability of the  $k$ th response class is

$$p_{ij \cdot k} = w_{i \cdot k} + x_i y_{j \cdot k}.$$

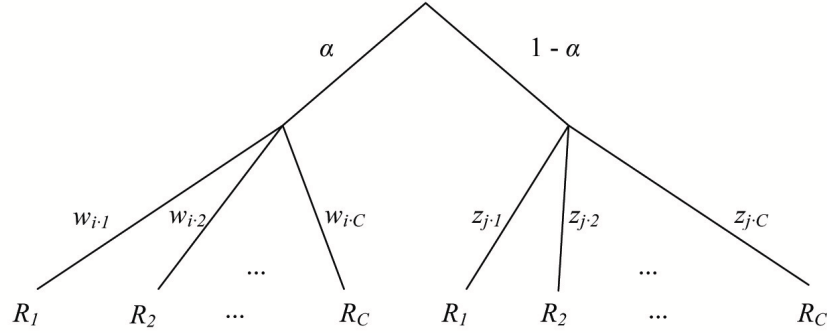


Figure 9: A standard  $C$ -class processing tree for unordered processing

In this tree, the two factors interact.

The two standard trees of Schweickert and Chen (2008) are special cases of the standard tree considered here in Figure 4 and Equation (6). This can be easily seen from Figure 5. If  $x_i = 0$ , the tree illustrated in Figure 5 can be written equivalently as the standard  $C$ -class tree for unordered processes illustrated in Figure 9. If  $b = 0$ , the arc labeled  $b$  and its descendents can be deleted from the tree in Figure 5. The remaining tree can be written equivalently as the standard  $C$ -class tree for ordered processes.

Through Theorem 1 a processing tree or rate tree can be constructed from experimental data, or the data will show that no tree is possible in which factors selectively influence processes and produce multiplicative interactions. The proof of Theorem 1 gives a way of calculating model parameters from a probability or rate matrix stack. Theorem 2 shows that parameters for a model are not unique, but given two sets of valid parameter values, values in one set can be obtained through transformations of those in the other set. Theorem 3 is potentially very informative, because it can be used to establish the order in which two mental processes are executed. If two different orders for processes influenced by the factors are possible, the data must have a special form. The subject may actually use a quite complicated tree to perform a task, but according to Theorem 4, the standard processing or rate tree in Figure 4 (or Figure 7) will suffice, if certain conditions are met. Finally, different subjects may use different

trees, or the same subject may use different trees on different trials, and for some of these trees certain responses may not be possible. Nonetheless, from Theorem 5, the standard tree suffices for the mixture of trees if certain conditions are met.



### **Author Note**

Richard Schweickert, Department of Psychological Sciences, Purdue University; Zhuangzhuang Xi, Department of Psychological Sciences, Purdue University, United States.

This work is based on the master's thesis of the second author supervised by the first author. We thank Ehtibar Dzhafarov and Zygmunt Pizlo for helpful discussions. We are grateful to the reviewers for exceptionally conscientious and useful work. Portions of this work were sponsored by AFOSR Grant FA9550-09-1-0252 to Richard Schweickert and Ehtibar Dzhafarov.

Richard Schweickert, Department of Psychological Sciences, 365A Peirce Hall, Purdue University, 703 Third Street, Purdue University, West Lafayette, Indiana 47907, United States; swike@psych.purdue.edu.

## References

- Batchelder, W. H., & Riefer, D. M. (1990). Multinomial processing models of source monitoring. *Psychological Review*, 97, 548–564.
- Batchelder, W. H., & Riefer, D. M. (1999). Theoretical and empirical review of multinomial process tree modeling. *Psychonomic Bulletin & Review*, 6, 57–86.
- Bishara, A. J., & Payne, B. K. (2009). Multinomial process tree models of control and automaticity in weapon misidentification. *Journal of Experimental Social Psychology*, 45, 524–534.
- Buchner, A., & Erdfelder, E. (2005). Word frequency of irrelevant speech distractors affects serial recall. *Memory & Cognition*, 33, 86–97.
- Clark, F. C. (1958). The effect of deprivation and frequency of reinforcement on variable-interval responding. *Journal of the Experimental Analysis of Behavior*, 1, 221–228.
- Conrey, F. R., Sherman, J. W., Gawronski, B., Hugenberg, K., & Groom, C. J. (2005). Separating multiple processes in implicit social cognition: The quad model of implicit task performance. *Journal of Personality and Social Psychology*, 89, 469–487.
- De Houwer, J. (2003). The extrinsic affective simon task. *Experimental Psychology*, 50, 77–85.
- De Houwer, J., & Eelen, P. (1998). An affective variant of the simon paradigm. *Cognition & Emotion*, 12, 45–62.
- Dodson, C. S., Holland, P. W., & Shimamura, A. E. (1998). On the recollection of specific- and partial-source information. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 24, 1121–1136.
- Dzhafarov, E. (2003). Selective influence through conditional independence. *Psychometrika*, 68, 7–25.

- Erdfelder, E., Auer, T.-S., Hilbig, B. E., Assfalg, A., Moshagen, M., & Nadarevic, L. (2009). Multinomial processing tree models: A review of the literature. *Zeitschrift für Psychologie/Journal of Psychology*, 217, 108–124.
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, 74, 1464–1480.
- Hu, X., & Batchelder, W. H. (1994). The statistical analysis of general processing tree models with the em algorithm. *Psychometrika*, 59, 21–47.
- Jacoby, L. L., Begg, I. M., & Toth, J. P. (1997). In defense of functional independence: Violations of assumptions underlying the PDP? *Journal of Experimental Psychology: Learning, Memory & Cognition*, 24, 3–26.
- Kujala, J. V., & Dzharfarov, E. N. (2008). Testing for selectivity in the dependence of random variables on external factors. *Journal of Mathematical Psychology*, 52, 128–144.
- Li, X., Schweickert, R., & Gandour, J. (2000). The phonological similarity effect in immediate recall: Positions of shared phonemes. *Memory & Cognition*, 28, 1116–1125.
- Meiser, T. (2005). A hierarchy of multinomial models for multidimensional source monitoring. *Methodology*, 1, 2–17.
- Meiser, T., & Bröder, A. (2002). Memory for multidimensional source information. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 28, 116–137.
- Payne, B. K. (2001). Prejudice and perception: The role of automatic and controlled processes in misperceiving a weapon. *Journal of Personality and Social Psychology*, 81, 181–192.
- Riefer, D., & Batchelder, W. H. (1995). A multinomial modeling analysis of the recognition-failure paradigm. *Memory & Cognition*, 23, 611–630.

- Roberts, S. (1987). Evidence for distinct serial processes in animals: The multiplicative-factors method. *Animal Learning & Behavior*, 15, 135–173.
- Roodenrys, S., & Miller, L. M. (2008). A constrained rasch model of trace redintegration in serial recall. *Memory & Cognition*, 36, 578–587.
- Schweickert, R. (1993). A multinomial processing tree model for degradation and redintegration in immediate recall. *Memory & Cognition*, 21, 168–175.
- Schweickert, R., & Chen, S. (2008). Tree inference with factors selectively influencing processes in a processing tree. *Journal of Mathematical Psychology*, 52, 158–183.
- Stahl, C., & Degner, J. (2007). Assessing automatic activation of valence: A multinomial model of east performance. *Experimental Psychology*, 54, 99–112.
- Thomas, R. D. (2006). Processing time predictions of current models of perception in the classic additive factors paradigm. *Journal of Mathematical Psychology*, 50, 441–455.
- Thorn, A. S. C., Gathercole, S. E., & Frankish, C. R. (2005). Redintegration and the benefits of long-term knowledge in verbal short-term memory: An evaluation of schweickert's (1993) multinomial processing tree model. *Cognitive Psychology*, 50, 133–158.
- Thorn, A. S. C., & Page, M. (2009). *Interactions Between Short-Term and Long-Term Memory in the Verbal Domain*. Hove, East Sussex, UK: Psychology Press.

Table 1

*Change of Parameter Bounds*

---


$$\text{If } c > 0, \quad b_k c \geq -\min_j \{y_{j \cdot k}\} \quad (10)$$

$$\text{If } c < 0, \quad b_k c \leq -\max_j \{y_{j \cdot k}\} \quad (11)$$

$$b_k c \geq \max_{x_i \neq x_1} \left\{ \frac{w_{i \cdot k} - w_{1 \cdot k} + w_{1 \cdot k}^* - 1}{x_i - x_1} \right\} \quad (12)$$

$$b_k c \leq \min_{x_i \neq x_1} \left\{ \frac{w_{i \cdot k} - w_{1 \cdot k} + w_{1 \cdot k}^*}{x_i - x_1} \right\} \quad (13)$$

$$x_1^* (b_k + \frac{1}{c} \min_j \{y_{j \cdot k}\}) \geq \max_j \{z_{j \cdot k} + x_1 y_{j \cdot k}\} + w_{1 \cdot k} - w_{1 \cdot k}^* - 1 \quad (14)$$

$$x_1^* (b_k + \frac{1}{c} \max_j \{y_{j \cdot k}\}) \leq \min_j \{z_{j \cdot k} + x_1 y_{j \cdot k}\} + (w_{1 \cdot k} - w_{1 \cdot k}^*) \quad (15)$$

$$\text{If } c > 0, \quad c(1 - b_k) \geq \max_j \{y_{j \cdot k}\} \quad (16)$$

$$\text{If } c < 0, \quad c(1 - b_k) \leq \min_j \{y_{j \cdot k}\} \quad (17)$$

$$-\frac{x_1^*}{x_I - x_1} \leq c \leq \frac{1 - x_1^*}{x_I - x_1} \quad (18)$$

$$w_{1 \cdot k}^* + \min_{x_i = x_1} \{w_{i \cdot k}\} - w_{1 \cdot k} \geq 0 \quad (19)$$

$$w_{1 \cdot k}^* + \max_{x_i = x_1} \{w_{i \cdot k}\} - w_{1 \cdot k} \leq 1 \quad (20)$$

$$\frac{1}{c} + \sum_{k=1}^C b_k = 1 \quad (21)$$


---

Table 2

*Change of Parameter Order Bounds I*

---


$$w'_{1 \cdot k} - w_{1 \cdot k} + z'_{j^*(k) \cdot k} - z_{j^*(k) \cdot k} - c_k x_{i^*(k)} y'_{j^*(k)} = 0 \quad (22)$$

$$\sum_k c_k = 0 \quad (23)$$

$$\sum_k c_k x_{i^*(k)} = -1 \quad (24)$$

$$c_k \geq 0 \text{ with } x_{i^*(k)} \leq \min_i \{x_i\}, \text{ or } c_k < 0 \text{ with } x_{i^*(k)} > \max_i \{x_i\} \quad (25)$$

$$c_k \geq 0 \text{ or } c_k \leq -\frac{\max_j \{y_{j \cdot k}\}}{y'_{j^*(k)}} \quad (26)$$

$$c_k \leq \frac{1}{y'_{j^*(k)}} \min_i \left\{ \frac{w_{i \cdot k} - w_{1 \cdot k} + w'_{1 \cdot k}}{x_i} \right\} \quad (27)$$

$$c_k \leq \frac{1}{\max_{x_i > x_{i^*(k)}} \{x_i\} - x_{i^*(k)}} \text{ or } c_k \geq \frac{1}{\min_{x_i < x_{i^*(k)}} \{x_i\} - x_{i^*(k)}} \quad (28)$$

$$c_k \leq 0 \text{ or } c_k \geq \frac{\max_j \{y_{j \cdot k}\}}{1 - y'_{j^*(k)}} \quad (29)$$

$$c_k \geq \frac{1}{y'_{j^*(k)}} \max_i \left\{ \frac{w_{i \cdot k} - w_{1 \cdot k} + w'_{1 \cdot k} - 1}{x_i} \right\} \quad (30)$$


---

Note: From the definition of  $y'_j$ , it follows that

$$\sum_k c_k y'_{j^*(k)} = -1.$$

Table 3

*Change of Parameter Order Bounds II*

---


$$\begin{aligned}
& c_k > 0 \text{ with } h(k) \geq -\frac{\min_i \{x(i, k)\}}{c_k}, \\
& \text{or } c_k < 0 \text{ with } h(k) \geq -\frac{\max_i \{x(i, k)\}}{c_k}, \text{ or } c_k = 0
\end{aligned} \tag{31}$$

$$\begin{aligned}
& c_k > 0 \text{ with } h(k) \leq 1 - \frac{\max_i \{x(i, k)\}}{c_k}, \\
& \text{or } c_k < 0 \text{ with } h(k) \leq 1 - \frac{\min_i \{x(i, k)\}}{c_k}, \text{ or } c_k = 0
\end{aligned} \tag{32}$$

$$\begin{aligned}
& c_k > 0 \text{ with } g(k) \leq \min_j \{y(j)\}, \\
& \text{or } c_k < 0 \text{ with } g(k) \geq \max_j \{y(j)\}, \text{ or } c_k = 0
\end{aligned} \tag{33}$$

$$\begin{aligned}
& c_k > 0 \text{ with } g(k) \geq \max_j \{y(j)\} - \frac{1}{c_k}, \\
& \text{or } c_k < 0 \text{ with } g(k) \leq \min_j \{y(j)\} - \frac{1}{c_k}, \text{ or } c_k = 0
\end{aligned} \tag{34}$$

$$0 \leq y(j) \leq 1 \tag{35}$$

$$0 \leq x(i, k) \leq 1 \tag{36}$$

$$\sum_{k=1}^C c_k = 0 \tag{37}$$

$$\sum_{k=1}^C c_k h(k) = -1 \tag{38}$$

$$\sum_{k=1}^C c_k g(k) = -1 \tag{39}$$

$$x(i, k)y(j) \leq p_{ij \cdot k} \tag{40}$$


---

Note: A restriction not required for one parameter may be required for another.

Table 4

*Modifications on Theorems 3 for Generalized Rate Models*

Parameters Representing Rates	Boundaries Not Required in Statement (2)	Restrictions Not Required in Statement (3)
$x_i$	NA	(32)
$y_{j \cdot k}$	NA	(34), (37), (39)
$w_{i \cdot k}$	NA	NA
$z_{j \cdot k}$	NA	NA
$y'_j$	(29)	upper bound of (35)
$x'_{j \cdot k}$	(28)	upper bound of (36), (37), (38)
$w'_{i \cdot k}$	(30)	NA
$z'_{j \cdot k}$	NA	NA